Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Biología Molecular

Role of pluripotency factor NANOG in stratified epithelia and cancer

Doctoral Thesis

Adelaida Rosa Palla Cachorro

Madrid, 2014

Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Biología Molecular

Role of pluripotency factor NANOG in stratified epithelia and cancer

Adelaida Rosa Palla Cachorro
BSc, Biotechnology

The entirety of the work presented in this Thesis has been carried out at the Tumor Suppression Group in the Spanish National Cancer Research Centre (CNIO, Madrid), under the supervision and direction of

Dr. Manuel Serrano Marugán

Madrid, 2014

cnio
Centro Nacional
de Investigaciones
Oncológicas

Dr. Manuel Serrano Marugán, Director of the Molecular Oncology Program and Head of the Tumor Suppression group (CNIO)

CERTIFIES:

That the Doctoral Thesis entitled "**ROLE OF PLURIPOTENCY FACTOR NANOG IN STRATIFIED EPITHELIA AND CANCER**" developed by **Mrs. Adelaida Rosa Palla Cachorro, BSc, MSc** meets all requirements to obtain the degree of **Doctor of Philosophy (PhD) in Molecular Biosciences**, that will, with the aforementioned objective, be defended at the Universidad Autónoma de Madrid. This thesis has been carried out under my supervision and I authorize it to be presented to the Thesis Tribunal accordingly.

I hereby issue this certification in Madrid on April the 29th, 2014.

Manuel Serrano, PhD
Director of the Molecular Oncology Program
Tumor Suppression Group Head

Spanish National Cancer Research Centre (CNIO)
3 Melchor Fernandez Almagro street
Madrid 28029, SPAIN

A mis padres y a mi hermano

ACKNOWLEDGEMENTS

La realización de esta tesis no hubiese sido posible sin el apoyo de muchas personas que han hecho de esta experiencia única e inolvidable.

En primer lugar, debo agradecer a Manolo, quien me ha dado la oportunidad de formarme en su laboratorio y gracias a infinitas horas de reunión en su despacho me ha enseñado a tener pensamiento y rigor científico, me ha animado durante momentos de agotamiento, y a quien indudablemente siempre tendré de referencia en el ámbito profesional, y es mi "role model" a seguir en mi futuro profesional como científica.

A mi familia. Mamá y Papá por todo su sacrificio es que estoy aquí hoy, y por siempre creer ciegamente en mí y pujarme a hacer lo que me apasiona cueste lo que cueste. Muchas gracias por darme la oportunidad de venir aquí y por toda la formación que me han dado. A mi hermano Robi, no hubiese sido posible sin ti, por siempre apoyarme, por tus risas, por también siempre creer en mi, y estar allí todos los días preocupándote por todos mis problemas del laboratorio y por tus consejos. A mi hermana María José, que también a la distancia me ha dado consejos y a todos mis sobrinos, Serena, Isabela, Walter, y ahora ¡un nuevo miembro de la familia que está por venir!

A mi Dani, que día tras día me ha dado fuerzas y consejos para seguir adelante, y por esperarme tantas noches cuando llego tan tarde siempre con una sonrisa y animándome, y aún así ¡querer pasar todo un futuro juntos!

Quería también agradecer a Daniela que me ha ayudado en mis inicios en la ciencia, me ha dado la mejor discusión científica y ha tenido paciencia conmigo para enseñarme como hacer ciencia de élite y de la mejor manera posible. Gracias por tus consejos y apoyo, por las largas horas de conversación en la sala de patología, y también por tu mentorship. Juntas hemos tenido que pasar las dificultades que ha traído el gran proyecto de Nanog, pero que me han hecho madurar mucho en el ámbito profesional y de la ciencia. ¡Esperemos que pronto podamos celebrar nuestros papers!

El ambiente del laboratorio durante todos estos años ha sido una experiencia inolvidable y no solo me llevo un doctorado, sino que me llevo amigos que no tengo duda de que me acompañarán toda la vida.

Quiero especialmente agradecer a mi predoc team que aunque ninguno nos parezcamos en gustos y parezcamos incompatibles hemos pasado los mejores momentos juntos: Helencita, ¡ya te extraño, imagínate!, sólo espero que hagas el postdoc cerca de mí, Ianirita, por tu cariño y apoyo y ¡mucho ánimo que si llegamos juntas salimos juntas!, Lluc, ¡a por tu plan A que todas

te seguimos!, Meri G, comenzamos juntas, eres única e inolvidable, Lucy, chocamos en todos nuestras ideologías pero he disfrutado mucho de nuestros diferentes puntos de vista, no cambies nunca, ¡eres inigualable! y Miguel, por tu positividad contagiosa durante la tesis. Incluyo también a Meri A que me ayudó en mis comienzos, me bautizó princesita, por sus consejos y guía y que tengo en mis mejores recuerdos. A todos por acompañarme no solo en mis momentos buenos sino también para superar los malos, por compartir los mejores viajes juntos, por el apoyo mutuo en el laboratorio, por los jueves en Montoya, por hacer el día a día divertido en el labo, por las miles de horas llevaderas juntos en cultivos y patología, y hasta por momentos de encierre en barrera.

A Ana, por su buena voluntad y ayuda en el labo y sus consejos, ¡seguro que nos veremos en EEUU otra vez! To Han, your advice and support to keep going have been priceless, I'm sure you'll get to the top! A Cristina, que es uno de los pilares del laboratorio y por su paciencia a quien puedes recurrir para cualquier problema. A Susana porque no hay otra igual y por el apoyo en múltiples técnicas del labo. A Martina que ha hecho del tubo unforgettable y por su cariño y ánimo en todo momento. To Cian, my benchmate, for hours and hours of discussing science, late nights in the tubo, and also for your personal advice and support. To Tim, my deskmate, who also has given me a lot of scientific discussion and I have always appreciated your point of view. A Ralph, con tus puntos de vista tan controversiales que han hecho las horas en la poyata bastante divertidas. A Fabian por darle al ambiente del laboratorio un aire tan cálido y por sus consejos. A Ben, mi compañero del tubo por su personalidad que hizo del tubo también un sitio inolvidable. A mis compañeras latinas, Nora, Marinela y Águeda por comprenderme a veces mejor que otros y por sus conversaciones. A Gina por su compañía en tantos fines de semana y noches y por su infinita alegría. A Merche, por su visón maternal y sus conversaciones, y como no, por todos los consejos de boda. A los chicos del labo, Dani, Christian y Juanma, con los que también me reído mucho en el labo y por los momentos Montoya.

A nuestros nuevos predocs que tienen el futuro de la tesis por delante y que les deseo mucho éxito: Iole, Aksinya, Noelia, Raquel y Miguel Ángel.

De los supresores y telómeros estoy agradecida a todos, presentes y pasados, por su apoyo y ayuda en el labo, y discusión científica, Rosa, Pablo, Nani, Paula, Isabel, Elisa, Arantza, Elsa, Sandrina, Bruno, Antonio.

A mi comité de tesis, Miguel, Paco y Mirna por dar su granito de arena en mi proyecto y tener una visión externa y consejos para mi tesis.

A todos nuestros colaboradores, que por su ayuda hemos podido realizar todas nuestras publicaciones.

No menos importante quería agradecer a todas las personas que aunque no estén en el laboratorio ni en la ciencia han sido indispensable para darme todo el apoyo necesario para llevar a cabo mi doctorado. A José Ignacio y a toda su familia que han estado de apoyo durante mis primeros años de tesis. A mis amigas de la uni, entre ellas Ana por sus salidas y conversaciones para desconectar del laboratorio. A las chicas del máster, Laia y Patri Nieto que disfrutamos un año intenso y después muchas quedadas durante la tesis. A mis amigas de Venezuela, siempre pienso en ustedes aunque estén lejos y por apoyarme en momentos difíciles, entre ellas Patricia con la que más he compartido en Madrid y por enseñarme que ¡hay un mundo fuera del laboratorio y de la ciencia!

No puedo olvidarme de todas las unidades del CNIO, que sin su ayuda técnica ni la mitad de los resultados hubiese sido posible, a todas las chicas de Patología Comparada entre ellas Marta Cañamero, a Genómica, con la ayuda de Orlando para entender todas las variantes de Nanog, a Citometría de Flujo con la ayuda de Lola y a Confocal con la ayuda de Diego. A los bioinformáticos, Osvaldo y Gonzalo, que gracias a ellos le hemos podido dar sentido a nuestros resultados. En el animalario también quería agradecer a Gema que me ha dado mucha ayuda con los ratones y por hacer llevaderas las horas en barrera.

¡Muchas gracias a TODOS!

INDEX

SUMMARY

NANOG is a master transcription factor critical for the acquisition of both embryonic and induced pluripotency. Here we have further analyzed NANOG's function in adult tissues and cancer. First we have uncovered NANOG expression in adult tissues, where it is restricted to the basal layer of epithelia. To determine its function we have used genetically modified mice with gain-of-function (inducible *Nanog* transgenic mice). In adult tissues, we have discovered that ubiquitous NANOG overexpression promotes hyperplasia of specific epithelia such as esophagus and forestomach, and transcriptionally activates key mitotic mediators selectively in stratified epithelia, but not in other tissues. Specifically, we have observed that NANOG can directly bind and regulate the mitotic *Aurka* gene in epithelial cells. The mitotic function of NANOG is also maintained in cancers that are derived from stratified epithelia, specifically, squamous cell carcinomas. Therefore, we analyzed the oncogenic effect of NANOG *in vivo* in squamous cell carcinomas. In a tumor resistant mouse background (C57BL/6) we have found *Nanog* overexpression in epithelia promotes squamous cell carcinoma by directly activating epithelial-mesenchymal transition (EMT) and stemness genes. These results unravel NANOG as an oncogenic driver in skin tumors and strengthen the idea of pluripotency transcription factors as tumorigenic in a cancer context. NANOG is gaining relevance in cancer since its expression has been detected in a large panel of tumors and cancer cell lines, but its presence in human cells is confounded by the presence of multiple and highly similar paralogs. In particular, there are three paralogs encoding functional proteins, namely, *NANOG1*, *NANOG2* and *NANOGP8*, and at least eight additional paralogs that do not encode full-length NANOG proteins. We have examined *NANOG* family expression using a multi-*NANOG* PCR that amplifies the three functional paralogs and most of the non-functional ones. As anticipated, we found that human embryonic stem cells (ESCs) express large amounts of *NANOG1* and, interestingly, they also express *NANOG2*. In contrast, most human cancer cells tested express *NANOGP8* and the non-coding paralogs *NANOGP4* and *NANOGP5*. Notably, in some cancer cell lines, the NANOG protein levels produced by *NANOGP8* are comparable to those expressed by *NANOG1* in pluripotent cells. NANOGP8 presents 2 aminoacid changes in important domains of the protein, therefore we wanted to address if it maintains its pluripotent activity during reprogramming of differentiated cells to induced pluripotency. Our study demonstrates that NANOGP8 is as active as NANOG1 in the reprogramming of human and murine fibroblasts into induced pluripotent stem cells (iPSCs). These results demonstrate that cancer-associated NANOGP8 can contribute to promote de-differentiation and/or cellular plasticity. Altogether, we conclude that NANOG is functional in specific adult tissues and its overexpression is relevant in cancer.

RESUMEN

NANOG es un factor de transcripción crítico para la adquisición de pluripotencia embrionaria e inducida. En este trabajo nos hemos enfocado en el papel de NANOG en tejidos adultos y en cáncer. En primer lugar, hemos descubierto la expresión de NANOG en tejidos adultos, específicamente en la capa basal del epitelio estratificado. En segundo lugar, utilizando ratones modificados genéticamente, hemos descubierto que la sobreexpresión de NANOG activa mediadores de mitosis claves que promueven hiperplasia de muchos epitelios, incluyendo el estómago no glandular y el esófago. Esta función mitótica de NANOG también se mantiene en aquellos cánceres que derivan del epitelio estratificado, específicamente, carcinomas de células estratificadas. Adicionalmente, hemos analizado el efecto oncogénico de NANOG *in vivo* en este tipo de carcinomas. En una cepa de ratón resistente a la formación de tumores (C57BL/6), hemos encontrado que sobreexpresión de *Nanog* específicamente en epitelios promueve la progresión a carcinomas de células estratificadas in vivo, a través de la activación de genes relacionados a transición epitelio-mesenquimal y a desdiferenciación, o troncalidad. Estos resultados demuestran que NANOG posee capacidad tumorigénica en tumores cutáneos, reforzando la idea de que factores de transcripción relacionados a pluripotencia son oncogénicos. En línea con otros estudios, hemos confirmado expresión de NANOG en un gran número de tumores y líneas celulares humanas que le dan relevancia a NANOG como marcador tumoral. En humanos, existen un gran número de parálogos de *NANOG* con un alto grado de similitud que dificulta la detección de NANOG en muestras humanas. Concretamente, existen tres genes parálogos que codifican proteínas funcionales, *NANOG1*, *NANOG2* y *NANOGP8*, aparte de estos, también existen otra serie de parálogos que no codifican proteína al no poseer un marco abierto de lectura. Hemos examinado la expresión de los miembros de la familia *NANOG* utilizando una estrategia de PCR "multi-*NANOG*" que es capaz de amplificar todos los páralogos que dan lugar a proteína, así como a la mayoría de los que no. Hemos encontrado que células madre embrionarias expresan gran cantidad de *NANOG1*, y también de *NANOG2*. En contraste, la mayoría de las líneas celulares de cáncer humanas analizadas expresaban NANOGP8, y los parálogos no codificantes *NANOGP4* y *NANOGP5*. Hemos comprobado que la expresión de proteína derivada de *NANOGP8* en líneas tumorales es comparable a la de *NANOG1* en líneas pluripotentes. NANOGP8 presenta dos cambios de aminoácidos en dominios importantes de la proteína, por lo tanto queríamos saber si mantiene su actividad pluripotente durante reprogramación de células diferenciadas a pluripotencia inducida. Nuestro estudio demuestra que NANOGP8 es tan activo como NANOG1 en la reprogramación de fibroblastos humanos y de ratón a células madre pluripotentes inducidas. Estos resultados demuestran que NANOGP8 asociado al cáncer puede contribuir a promover desdiferenciación y/o plasticidad celular. En conclusión, hemos demostrado que NANOG es funcional en tejidos adultos específicos y su sobreexpresión es relevante en el cáncer.

KEY ABBREVIATIONS

- ALDH: aldehyde dehydrogenase
- AP: alkaline phosphatase
- Aurka: aurora kinase A
- Aurkb: aurora kinase B
- ALT: alanine transaminase
- BCC: basal cell carcinoma
- bp: base pair
- CDC20: cell division cycle 20
- cDNA: complementary DNA
- CDX2: caudal-type homeobox protein 2
- ChIP: chromatin immunoprecipitation
- CK: cytokeratin
- CoIP: co-immunoprecipitation
- CSC: cancer stem cell
- cSCC: cutaneous squamous cell carcinoma
- CXCR: CXC chemokine receptor
- DMBA: 7,12-dimethylbenz(a)anthracene
- DMEM: Dulbecco's modified Eagle's medium
- DMR: differentially methylated region
- DNA: deoxyribonucleic acid
- DOX: doxycycline
- ECL: enhanced chemiluminescence
- EMT: epithelial-mesenchymal transition
- Erk: extracellular signal-related
- ESCC: esophageal squamous cell carcinoma
- ESCs: embryonic stem cells
- Esrrb: estrogen-related receptor beta
- Fgf: fibroblast growth factor
- GPCR: G-protein-coupled receptors
- GSEA: gene set enrichment analysis
- Gsk3: glycogen synthase kinase 3
- H&E: hematoxylin and eosin
- HDAC: histone deactelyase
- HMG: high mobility group
- HNSCC: head and neck squamous cell carcinoma
- HPV: human papilloma virus
- HSC: hematopoietic stem cell
- ICM: inner cell mass
- IFE: interfollicular epidermis
- IHC: immunohistochemistry
- iPSCs: induced pluripotent stem cells
- ISH: *in situ* hybridization
- JAK: Janus-associated tyrosine kinase
- K5: cytokeratin 5
- Kif20a: kinesin family member 20A
- KO: knock-out
- LIF: leukemia inhibitory factor

- MAPK: mitogen-activated protein kinase
- MEFs: Mouse embryonic fibroblasts
- MET: mesenchymal-epithelial transition
- miRNA: microRNA
- NLS: nuclear localization signal
- NuRD: nucleosome remodeling deacetylase
- OCT3/4: octamer-binding transcription factor 3/4
- ORF: open reading frame
- OSKM: Oct4, Sox2, Klf4 and cMyc
- PcG: polycomb group
- Pdgfra: platelet derived growth factor receptor alpha
- PGCs: primordial germ cells
- PI3K: phosphatidylinositol-3-OH kinase
- PITX2: pituitary homeobox 2
- PKC: protein kinase C
- Prrx1: paired related homeobox 1
- PTM: post-translational modification
- qRT-PCR: quantitative real-time PCR
- RNA: ribonucleic acid
- rtTA: reverse tetracycline transactivator
- SC: stem cell
- SCC: squamous cell carcinoma
- SDF-1: stromal cell-derived factor-1
- siRNA: small interfering RNA
- SpSCC: spindle cell carcinoma
- TBIL: total bilirubin
- TE: trophectoderm
- Tet: ten eleven translocation enzymes
- TNFα: tumor necrosis factor alpha
- TPA: 12-O-tetradecanoylphorbol-13-acetate
- UPS: ubiquitin-proteasome system
- WR: tryptophan-rich domain
- WT: wild-type

INTRODUCTION

1. Pluripotency-associated transcription factor NANOG

NANOG is a homeodomain-containing transcription factor, which was first discovered in embryonic stem cells (ESCs) for its capacity of maintaining the pluripotency and self-renewal of these cells (Chambers et al., 2003; Mitsui et al., 2003).

1.1. NANOG in embryonic development

Nanog is crucial during embryogenesis, as shown by the fact that *Nanog*-null embryos present early embryonic lethality (Mitsui et al., 2003). Specifically, it is expressed at two points during development, firstly for the correct specification of a pluripotent epiblast from the inner cell mass (ICM) at the blastocyst stage (Mitsui et al., 2003; Silva et al., 2009), and secondly it is also important for the development of primordial germ cells (PGCs) at E11.5 (Chambers et al., 2007; Yamaguchi et al., 2009) **(Figure 1)**.

In the ICM there are distinct subpopulations of cells. One of these subpopulations selectively expresses *Nanog* that then commits to become the epiblast, meanwhile the other subpopulation expresses GATA-binding factor 6 (GATA6) and commits into the primitive endoderm. The epiblast gives rise to the embryo proper maintaining pluripotency, while the primitive endoderm gives rise to extraembryonic structures (Chazaud et al., 2006) **(Figure 1)**.

Nanog is also crucial for PGC development, the precursors for oocyte and spermatozoa, which undergo extensive reprogramming (Surani et al., 2007). Nanog is expressed at E11.5 (Yamaguchi et al., 2005), where PGCs go through epigenetic remodeling including reactivation of the inactive X chromosome (female cells) and DNA demethylation (Surani et al., 2007) (**Figure 1)**; features shared by naïve pluripotent cells (Welling and Geijsen, 2013). Conditional knockdown of *Nanog* in migrating PGCs resulted in apoptosis of these cells (Yamaguchi et al., 2009), confirming its requirement for the specification of the germ cell lineage.

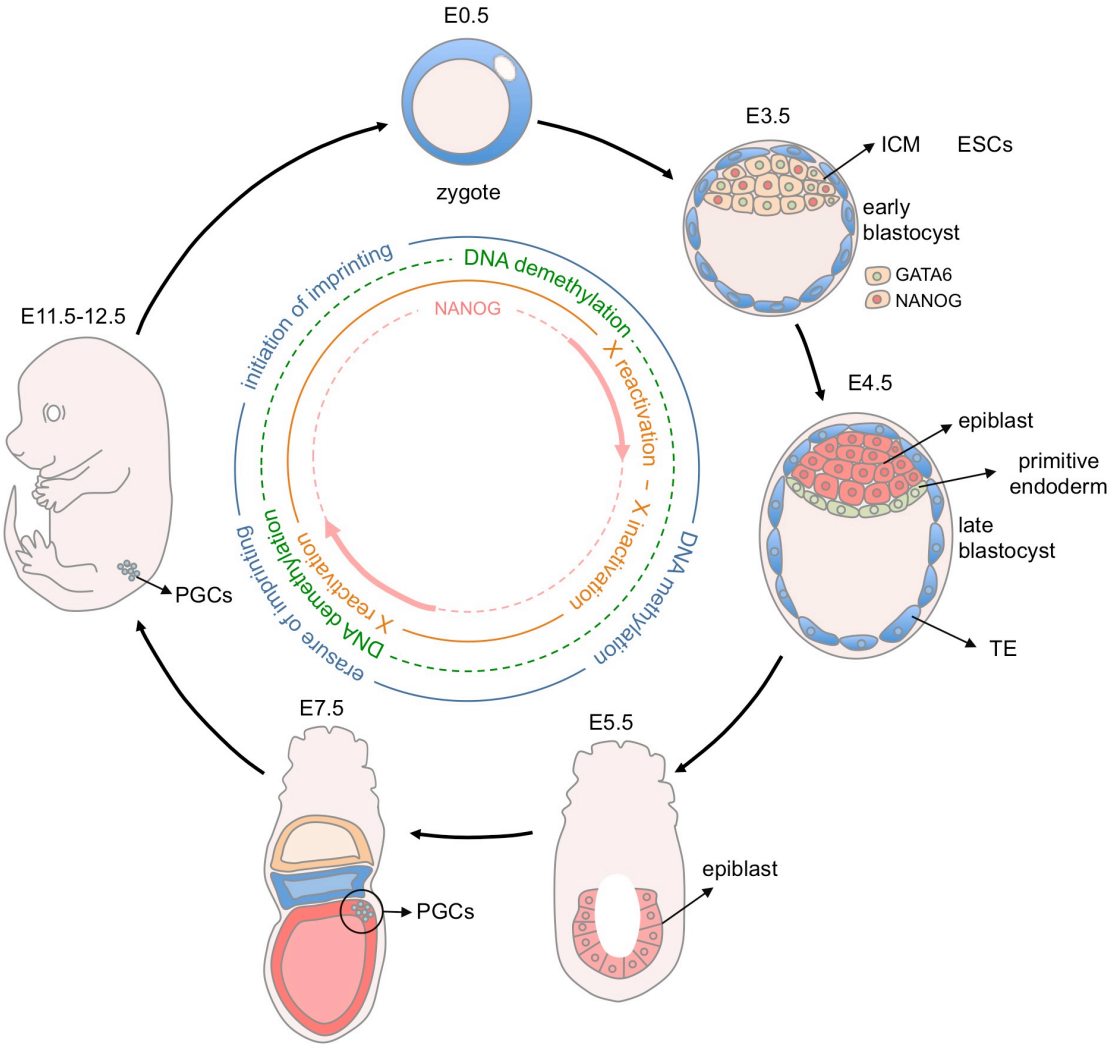


Figure 1. *Nanog* expression during embryonic development
The zygote is totipotent and maternally inherits many epigenetic modifiers and transcription factors, which promote development to the blastocyst stage. *Nanog* expression in the inner cell mass (ICM) of the early blastocyst (E3.5) defines the specification of the epiblast. Pluripotency in epiblast cells is shown by X reactivation and maintenance of demethylation. Once cells start to commit into different lineages, DNA methylation appears. Only the primordial germ cells directly derive from the epiblast and maintain their pluripotency where *Nanog* expression is required, associated to erasure of imprinting, DNA demethylation and X reactivation in these cells. Modified from (Chazaud et al., 2006).

1.2. Structure of NANOG

NANOG is a transcription factor with a single homeodomain that binds to DNA (Jauch et al., 2008). The *Nanog* gene encodes a 305-aminoacid protein, whose general structure consists of an N-terminal, a homeobox and a C-terminal domain (Chambers et al., 2003; Mitsui et al., 2003; Wang et al., 2003) **(Figure 2)**.

Nanog is conserved in the mammalian lineage, both placental and non-placental. Additionally it has shown to possess orthologs in eutherians, birds, axoltl and teleosts (Booth and Holland, 2004; Camp et al., 2009; Canon et al., 2006; Dixon et al., 2010; Schuff et al., 2012). No clear homologues appear in invertebrates (Booth and Holland, 2004), and *Nanog* seems to have evolved only in vertebrates (Theunissen et al., 2011). Most pluripotency genes are well

conserved between eutherian mammals (Theunissen et al., 2011) and although *Nanog* is central in this transcriptional network (Wang et al., 2006b), it is poorly conserved between species (only ~54% between mouse and human) (Chambers et al., 2003; Mitsui et al., 2003; Wang et al., 2003).

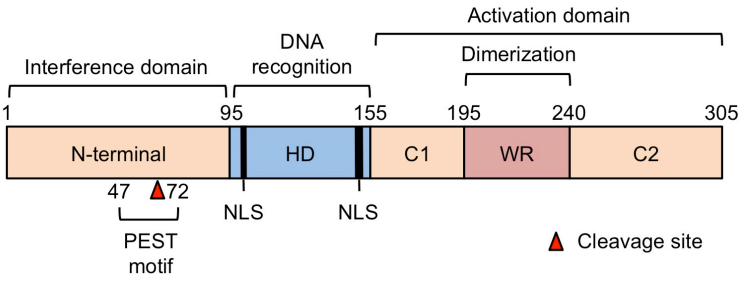


Figure 2. Structure of human NANOG
Summary of functional and structural domains of human NANOG. Red triangle shows cleavage site mediated by caspases. Black line shows NLS sites. Abbreviations: NLS, nuclear localization signal; HD, domeodomain; C1, C-terminal subdomain 1; WR, tryptophan-rich subodomain; C2, C-terminal subdomain 2.

1.2.1. Homeobox domain

The homeobox gene superfamily can be subdivided into classes, which contain numerous gene families. NANOG forms part of the ANTP class, and its closest sequence similarity is 50% with the NK2 family (Wang et al., 2003). The NANOG homeodomain comes out as a variant, since it lacks the N-terminal TN domain specific of the NK family, indicating it does not belong to any previously isolated subfamilies **(Figure 3)** (Harvey, 1996). The NANOG homeobox is the most conserved domain between its orthologs (85% identity between mouse and human) **(Figure 3)** (Canon et al., 2006; Chambers et al., 2003; Mitsui et al., 2003). The homeodomain determines the DNA binding specificity, and contains three alpha helices (helix I, helix II and helix III), where helix III is involved in DNA recognition (Gehring et al., 1994; Gruschus et al., 1997; Liu et al., 1990) **(Figure 3)**. The importance of this homeodomain has been further elucidated by a recent study, showing that the homeodomain of different Nanog orthologs is enough to establish pluripotency in *Nanog* knock-out (KO) mouse somatic cells (Theunissen et al., 2011).

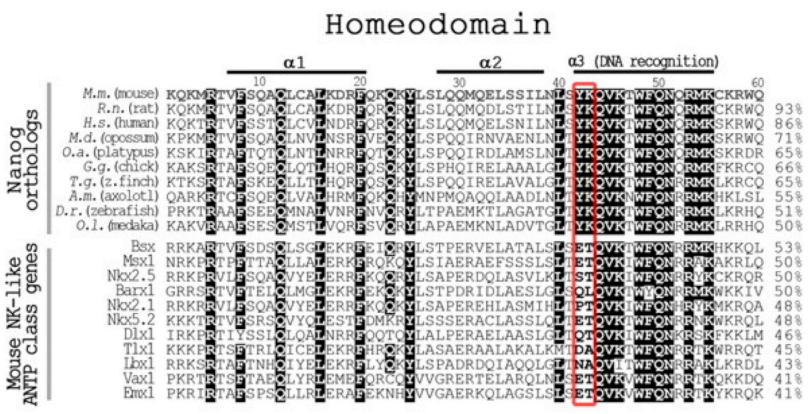


Figure 3. Homeodomain comparison between Nanog orthologues and NK-like class genes.
Sequence alignment of the homeodomains of Nanog orthologs and related mouse NK-like ANTP class proteins. Identical amino acids are highlighted in black. Red frame indicates residues unique to Nanog orthologs. Identity scores relative to mouse Nanog are shown on the right and were calculated using ClustalW2. Adapted from (Theunissen et al. 2011).

The consensus DNA sequence bound by NANOG has still to be clarified, and although different ChIP-seq studies have been carried out in embryonic stem cells (ESCs) (Boyer et al., 2005; Loh et al., 2006), the DNA binding sequence of NANOG still remains unclear.

Transcription factors require a nuclear import process to activate target genes, normally mediated by a nuclear localization signal (NLS) composed of short stretches of highly basic aminoacid residues (Kalderon et al., 1984). In human NANOG, two nuclear localization signals (NLS) have been identified within the homeodomain (Chang et al., 2009; Do et al., 2007) **(Figure 2)**.

1.2.2. Transactivation domains

Both the C-terminal and N-terminal regions of NANOG have transcription-related activities (Pan and Pei, 2003). The N-terminal region is serine rich, and appears to act as an interference domain, because its deletion leads to enhanced transactivation of target promoters **(Figure 2)** (Chang et al., 2009). The lack of the first serine residue was demonstrated to be essential for self-renewal of murine ESCs (Das et al., 2011). As described below, the N terminus is also important for many post-translational modifications that also have a direct effect on NANOG's function.

The C terminus can be divided into three subdomains, C1, tryptophan-rich (WR) subdomain, and C2 **(Figure 2)**. The C-terminal region has two potent transactivation domains, WR and C2, which mediate transcriptional activation (Do et al., 2009; Oh et al., 2005; Pan and Pei, 2005). Specifically, the glutamine motif in the WR subdomain and the acidic residues in CD2 are required for the transcriptional activation activity of the C terminus of human NANOG **(Figure 2)** (Do et al., 2009).

Additionally, the C terminus is key for NANOG's function since it is essential for homodimerization directly through its WR subdomain (Mullin et al., 2008; Wang et al., 2008). NANOG's dimerization is indispensable for pluripotency maintenance and self-renewal of ESCs (Mullin et al., 2008; Wang et al., 2008). Mutations of the tryptophan residues to alanine in the WR impede its dimerization and monomeric NANOG has impaired interaction to its known partners, Sall4, Zfp198, Dax1, Nac1 and Oct4 (Wang et al., 2008) **(Figure 2)**.

1.3. NANOG post-translational modifications

Pluripotency-associated transcription factors are highly expressed in pluripotent cells and become rapidly downregulated when these become differentiated. A number of different post-translational modifications (PTM) have been described that can regulate the levels of these

pluripotency factors to achieve an optimum balance between pluripotency and differentiation (Cai et al., 2012). To date, NANOG has been described to undergo phosphorylation, ubiquitylation, and caspase-mediated cleavage.

Phosphorylation

NANOG has been known to be a phosphoprotein for quite some time (Yates and Chambers, 2005), but only a recent study using mass spectrometry has demonstrated that human NANOG can be phosphorylated at multiple sites on its N terminus at proline-directed sites (Brumbaugh et al., 2014). Regarding its significance, one of the first functional studies described phosphorylation at multiple Ser/Thr-Pro motifs, which leads to recognition and binding of the prolyl isomerase Pin1 to NANOG. This interaction leads to stabilization of NANOG by preventing proteasome-mediated degradation. The presence of Pin1 and Pin1-Nanog interactions is important for mouse ESC self-renewal and teratoma formation (Moretto-Zita et al., 2010). An *in vitro* screening approach of a subset of kinases, revealed ERK2 and CDK1 could directly phosphorylate NANOG, therefore creating a potential connection between key signaling pathways and NANOG (Brumbaugh et al., 2014).

Ubiquitylation

The ubiquitin-proteasome system (UPS) is the main pathway responsible for eliminating intracellular proteins (Lecker et al., 2006). A PEST motif (rich in proline, glutamine, serine, and threonine) in the N terminus of human NANOG was recently described to undergo ubiquitylation. Deletion of this PEST motif increased the stability and half-life of NANOG (Ramakrishna et al., 2011) **(Figure 2)**.

Cleavage

Another type of PTM which can lead to rapid modifications of ESC function are mediated by site-specific proteases, not only known for their cell death program execution (Earnshaw et al., 1999), but also for other functions in specific contexts such as cell differentiation (Arama et al., 2003; Feinstein-Rotkopf and Arama, 2009). Upon differentiation of ESCs, caspase activity increased and specific sites in the N terminus and the homeodomain of mouse and human NANOG were identified **(Figure 2)**. Caspase-3 can directly mediate ESC differentiation by cleavage of NANOG, leading to its destabilization and subsequent degradation (Fujita et al., 2008).

1.4. NANOG in embryonic stem cells

ESCs are derived from the ICM of mammalian blastocysts. They are characterized by pluripotency, giving them capacity to differentiate into all cell types (Evans and Kaufman,

1981; Martin, 1981), as well as by self-renewal, which gives them the capacity to divide symmetrically and form two identical stem cell daughters (Burdon et al., 2002). One of the most important pathways through which mouse ESCs maintain their self-renewal in culture is signaling through the leukemia inhibitory factor (LIF) (Smith et al., 1988; Williams et al., 1988). LIF is usually provided to cultured ESCs as a recombinant protein or through a feeder layer of embryonic fibroblasts. LIF ultimately acts by activating STAT3, among others, through binding to LIF receptor and gp130 heterodimers that activate Janus-associated kinases (JAK) **(Figure 4)** (Burdon et al., 2002; Niwa et al., 1998). NANOG was first discovered for its ability to maintain mouse ESC self-renewal independently of the LIF/Stat3 pathway (Chambers et al., 2003; Mitsui et al., 2003). Specifically, the Jak-Stat3 pathway activates KLF4 that in turn activates SOX2. On the other hand, JAK also activates the phosphatidylinositol-3-OH kinase (PI3K)-Akt pathway, which regulates TBX3 expression that in turn stimulates NANOG. Both SOX2 and NANOG promote expression of the *Oct4* endogenous locus and therefore self-renewal of ESCs (Niwa et al., 2009) **(Figure 4)**.

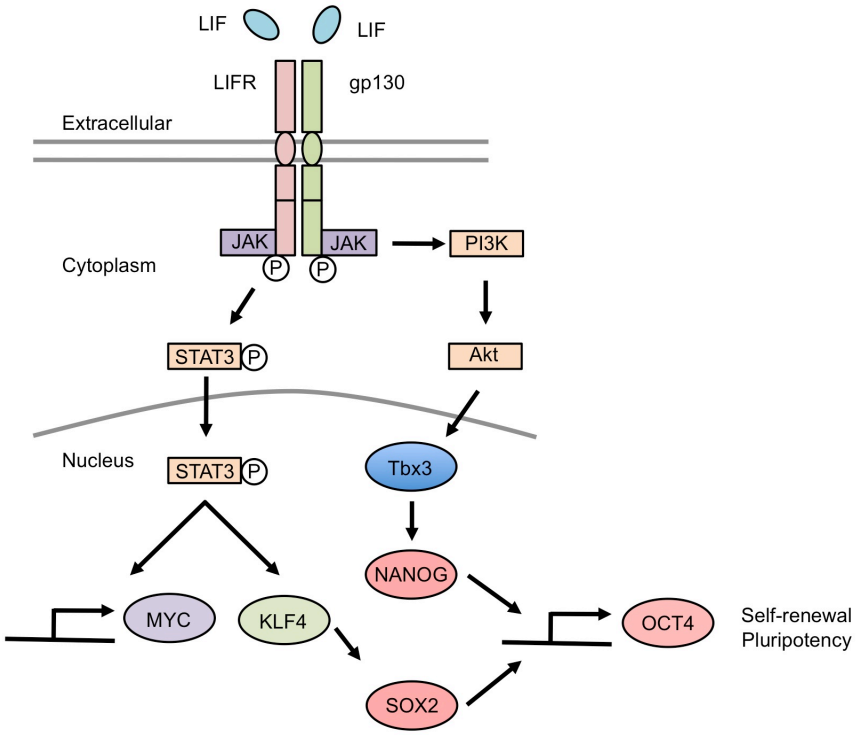


Figure 4. Embryonic self-renewal through the LIF signal pathway.
LIF binds to the heterodimer LIFR and gp130 activating JAK by phosphorylation. JAK activation phosphorylates Stat3, which translocates into the nucleus activating KLF4 and promoting expression cMYC. The PI3K/Akt pathway is also activated stimulating transcription of *Tbx3*. TBX3 and KLF4 activate *Nanog* and *Sox2,* respectively, which in turn maintain expression of *Oct3/4*.

Although NANOG is a key player in pluripotency of ESCs, studies have shown it does this in conjunction with other transcription factors, specifically OCT4 and SOX2. NANOG, together with SOX2 and OCT4 form the core transcriptional regulatory circuitry in ESCs (Boyer et al., 2005; Loh et al., 2006). OCT4 is a member of the POU class of homeodomain proteins that can

recognize an 8-bp DNA site through two DNA binding domains (POU_S and POU_{HD}), each contacting 4 bp in the major groove (Falkner and Zachau, 1984; Klemm and Pabo, 1996; Parslow et al., 1984; Phillips and Luisi, 2000). *Oct4* is expressed in pluripotent cells and specifies the ICM during embryonic development by counteracting differentiation into TE (Nichols et al., 1998). *Oct4* repression in mouse ESCs causes loss of self-renewal and differentiation into TE (Niwa et al., 2000) **(Figure 5)**. SOX2 is a member of the Sry-related high mobility group (HMG) box family of transcription factors that interacts with the DNA through the minor groove (Bowles et al., 2000). SOX2 is important for epiblast maintenance (Avilion et al., 2003) and *Sox2*-deficient mouse ESCs differentiate into TE (Ivanova et al., 2006; Masui et al., 2007).

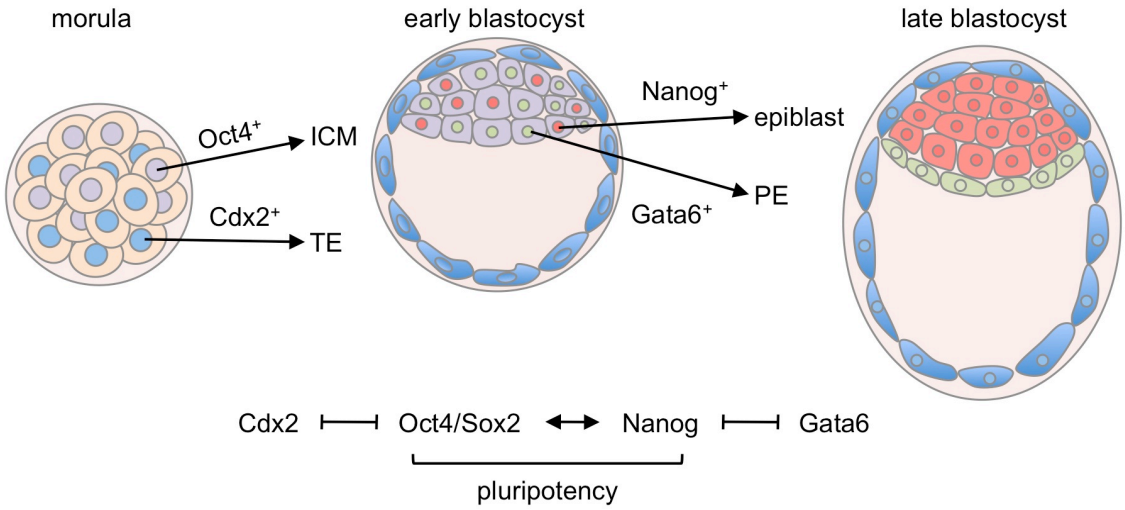


Figure 5. Lineage segregation in the blastocyst.
The morula stage at E3.0 shows distinct and mosaic expression of two transcription factors caudal-type homeobox protein 2 (CDX2) and octamer-binding transcription factor 3/4 (OCT3/4; also known as POU5F1) which decide the first cell fate to form the trophectoderm (TE) or the inner cell mass (ICM), respectively. Oct4 and Cdx2 regulate each other in a negative reciprocal manner to specify both lineages. The second cell fate decision occurs from the ICM, which segregates into the primitive endoderm (PE) and epiblast lineages. In this case Nanog positive cells specify into the epiblast that gives rise to the embryo proper, meanwhile Gata6 positive cells specify to the PE. Nanog and Gata6 also regulate their expression negatively. Modified from (Arnold and Robertson, 2009).

OCT4, SOX2 and NANOG cooperatively maintain the regulatory network responsible for self-renewal and pluripotency (Chambers and Tomlinson, 2009). OCT4, SOX2 and NANOG co-occupy many regulatory loci (Boyer et al., 2005; Loh et al., 2006), suggesting they may act coordinately to maintain the transcriptional program required for pluripotency. Additionally, studies have shown these core factors can cooperatively both activate and repress a number of target genes. Activation has been shown to involve recruitment of other transcriptional coactivators such as p300 when OCT4, SOX2 and NANOG bind together at active enhancers (Chen et al., 2008). On the other hand, silencing of developmental regulators is also important to maintain pluripotency, because expression of these developmental factors is associated with commitment to particular lineages This silencing appears to be mediated by Polycomb (PcG) group proteins, which are epigenetic regulators that facilitate maintenance of cell state through

gene silencing. Cobinding of OCT4, SOX2 and NANOG together with PcG proteins could mediate this silencing. (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006).

NANOG's role in pluripotency has been questioned by recent reports showing that *Nanog*-deficient ESCs can maintain their pluripotency. Two different concepts have emerged, establishment and maintenance of pluripotency. NANOG seems to be essential for the establishment of naïve pluripotency, clearly shown by (1) *Nanog* deletion fails in the formation of germ cells (Chambers et al., 2007), (2) the ICM of *Nanog*-deficient cells is unable to generate a pluripotent naïve epiblast (Messerschmidt and Kemler, 2010), and (3) a *Nanog*-deficient ICM cannot give rise to ESCs (Chambers and Tomlinson, 2009; Mitsui et al., 2003).

The main characteristics of ESCs with naïve or ground state pluripotency are that (1) they can form chimeras by reincorporating into the epiblast at the blastocyst stage and produce functional soma and germ cells, (2) both X chromosomes are activated in female cells, (3) they maintain self-renewal through the LIF/Stat3 pathway, (4) they commit when fibroblast growth factor (Fgf4)/Erk pathway is stimulated (Nichols and Smith, 2009). This last feature has led to the discovery that blockade of the MAPK/Erk pathway, as well as of glycogen synthase kinase-3 (Gsk3) with selective small molecule inhibitors (2i) is sufficient to stabilize and sustain ESCs with full pluripotency (Ying et al., 2008). Allelic regulation of *Nanog* could be key in regulation of naïve state pluripotency since cells in the naïve epiblast of the blastocyst as well as ESCs grown in 2i conditions switch from a monoallelic to a biallelic expression (Miyanari and Torres-Padilla, 2012). This observation of *Nanog*'s allelic regulation has been challenged by two recent publications and therefore should be further studied (Faddah et al., 2013; Filipczyk et al., 2013).

1.5. NANOG in reprogramming

In 2006, Takahashi and Yamanaka were able to reprogram differentiated somatic cells to an ESC-like state by ectopic expression of four transcription factors (*Oct4, Sox2, Klf4* and *cMyc*, abbreviated here as OSKM) (Takahashi and Yamanaka, 2006). These induced pluripotent stem cells (iPSCs) share the same properties with ESCs, including pluripotency and self-renewal. They rely on the same signaling pathways such as LIF and 2i addition, the core transcription factors (OCT4, SOX2 and NANOG) and share similar gene expression profiles (Hanna et al., 2010; Yamanaka, 2009).

Surprisingly, *Nanog* is not among the "Yamanaka" reprogramming factors and is dispensable at the initial steps of reprogramming. Nevertheless, *Nanog* has been shown to be essential at later steps of reprogramming (Faddah et al., 2013; Silva et al., 2008; Sridharan et al., 2009), where it is required to complete reprogramming to a "fully" reprogrammed state or what is known as the

naïve or ground pluripotent state (Chambers and Tomlinson, 2009). In the case of murine cell reprogramming, addition of mouse *Nanog* accelerates reprogramming with the "Yamanaka" cocktail (OSKM) (Hanna et al., 2009) and promotes the transition from pre-iPSCs to iPSCs. (Silva et al., 2009)

Nanog-deficient somatic cells cannot acquire a fully reprogrammed state and remain in a pre-iPS stage (Silva et al., 2009). It still remains unclear how *Nanog* plays this role during reprogramming. Some studies have shown that Nanog not only acts through direct regulation of genes involved in pluripotency, but also by physically interacting with chromatin remodellers such as NuRD (nucleosome remodeling deacetylase) and Tet1/2 (ten eleven translocation enzymes) (Festuccia et al., 2013). *Nanog* requirement during reprogramming can be bypassed by overexpressing NANOG target genes such as Esrrb together with 5'azacytidine treatment, a known DNA demethylating agent (Festuccia et al., 2012). Another study has shown that *Nanog* is dispensable during reprogramming by modifying the culture conditions. In detail, addition of ascorbic acid (vitamin C) to the medium permits *Nanog*-deficient cells to achieve complete reprogramming (Schwarz et al., 2014), this also is connected with epigenetic remodeling since ascorbic acid has been shown to directly stimulate Tet activity and induce DNA methylation in ESCs (Blaschke et al., 2013) **(Figure 6)**.

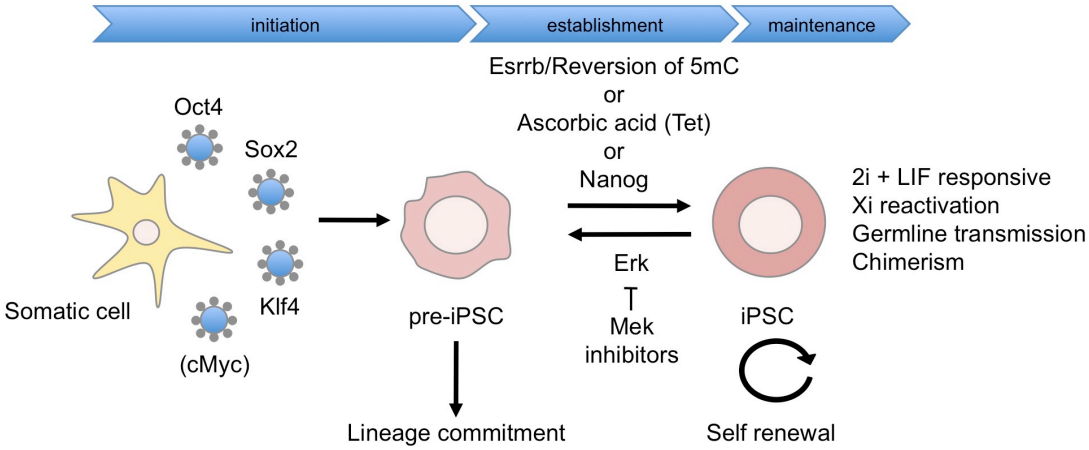


Figure 6. Requirement of NANOG during reprogramming.
Initiation of reprogramming starts with transduction of somatic differentiated cells, in this case retroviruses or lentiviruses containing the "Yamanaka" factors *Oct4, Klf4, Sox2* and *cMyc*. *cMyc* is not required in the reprogramming cocktail but increases the efficiency. This leads to appearance of cells in a transitional state known as pre-iPSCs, where there is expression of some pluripotency genes but these cells are prone to lineage commitment and have not acquired pluripotent features. *Nanog* is not required in this initial phase. In the establishment stage pre-iPSCs need to undergo complete reprogramming. *Nanog* expression is essential for this stage and it does this by inducing *Esrrb* expression and contributing to demethylation by recruiting Tet1. Nanog requirement can only be bypassed by overexpressing its target gene *Esrrb* used together with demethylating agent 5-azacytidine, or by adding to the medium ascorbic acid which is known to be a cofactor for Tet. iPSCs are 2i/LIF responsive, reactivate the inactive X chromosome, contribute to chimerism and can give rise to germline transmission. Once pluripotency is established, Nanog is no longer required in the maintenance phase and can be deleted without compromising self-renewal.

1.6. NANOG variants in the genome

1.6.1. Human variants

At least 11 *NANOG* paralogs have been identified in the human genome, which complicates its detection (Booth and Holland, 2004). The founding member of the family, *NANOG1 or* embryonic *NANOG*, is highly expressed in human embryonic stem cells (hESCs) (Chambers et al., 2003; Mitsui et al., 2003; Wang et al., 2003). In addition, *NANOG1* is an important component of the "Thomson" cocktail (*OCT4*, *SOX2*, *LIN28* and *NANOG1*) for the reprogramming of human cells into iPSCs (Yu et al., 2007). As mentioned before its homeodomain is functionally conserved among different species, demonstrated by a study where human *NANOG1* can replace mouse *Nanog* in the pre-iPSC to iPSC transition of murine cells (Theunissen et al., 2011). *NANOG2* is another paralog that is the product of a duplication event and lies on the same chromosome as *NANOG1*. It shares a similar intron/exon structure and a highly conserved promoter region. Recently *NANOG2* was found in human leukemic cells (Eberle et al., 2010). However, it is not known whether *NANOG2* is expressed in human embryonic stem cells (hESCs). In addition to *NANOG1* and *NANOG2*, there are 9 intronless paralogs, named *NANOGP2* to *NANOGP10*, under the control of promoter sequences unrelated to their parental *NANOG1* and *2* genes (Booth and Holland, 2004). *NANOGP8* is of particular interest because it is the only *NANOG* intronless paralog with an intact protein coding capacity and, indeed, its encoded protein only differs in 2 amino acids from NANOG1 (Zbinden et al., 2010). Of note, the two amino acids divergent in NANOGP8 relative to NANOG1 are located in important domains for transcriptional regulation (Chang et al., 2009; Do et al., 2009; Pan and Pei, 2005), opening the possibility that the function of NANOGP8 could differ from that of NANOG1 **(Figure 7)**. Interestingly, *NANOGP8* is expressed in many human cancer cell lines, where it has been reported to increase many properties associated with cancer stem cells, including clonogenicity and tumorigenicity (Ambady et al., 2010; Ishiguro et al., 2012; Jeter et al., 2009; Jeter et al., 2011; Uchino et al., 2012; Zbinden et al., 2010; Zhang et al., 2013; Zhang et al., 2010; Zhang et al., 2006a). However, there is no direct information on the activity of *NANOGP8* in ESCs or in pluripotency assays.

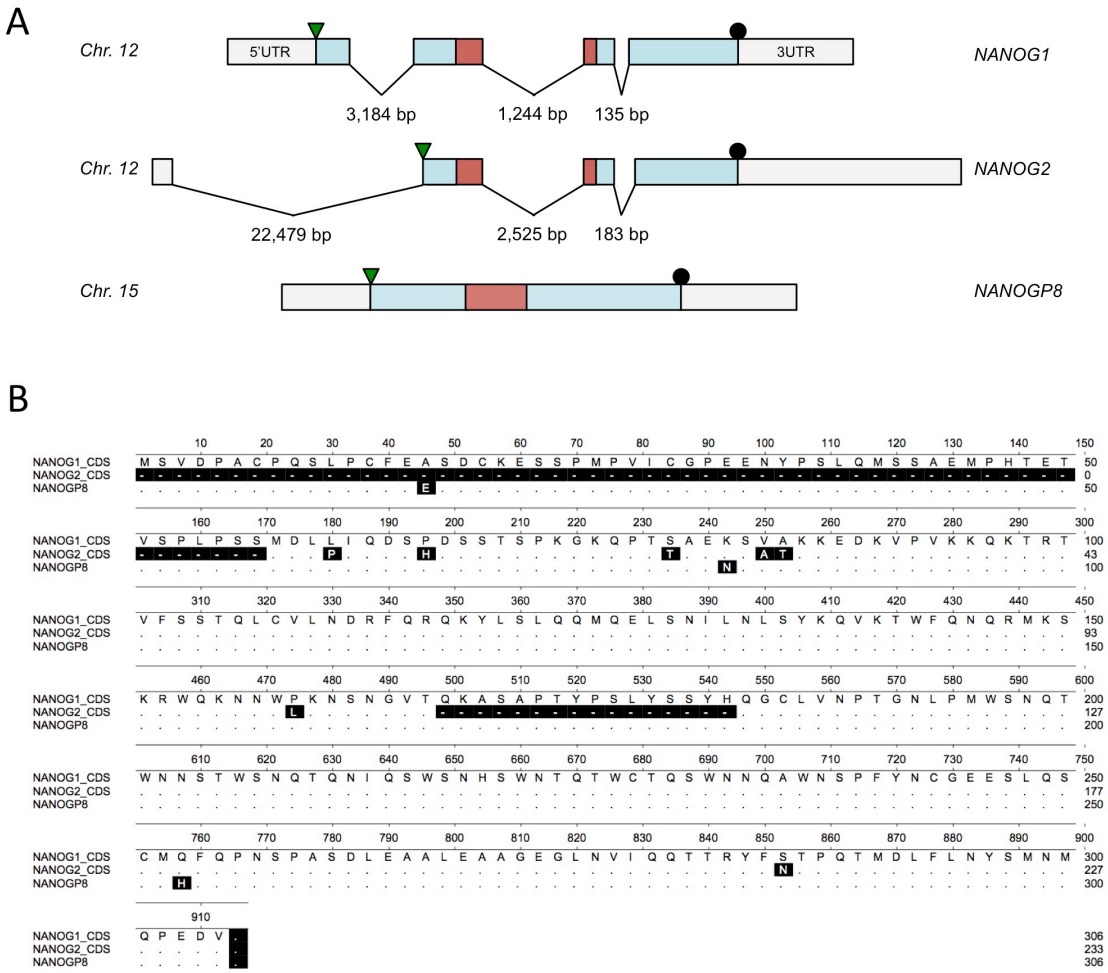


Figure 7. *NANOG* variants in the human genome

(A) Coding human NANOG variants gene structure. *NANOG1* and *NANOG2* have an exon/intron structure. Exons are represented by horizontal bars; 5' and 3' UTR are white; the coding region is blue and the homeodomain is colored in red. Introns are represented by lines and length in nucleotides is written underneath the introns. Green triangles indicate the start codon (ATG) and the black circle indicates the stop codon (TGA). *NANOGP8* is a retrogene that lacks introns. *NANOG1*, *NANOG2* and *NANOGP8* have a functional ORF and give rise to NANOG protein. Modified from (Booth and Holland, 2004).

(B) Alignment report comparing the protein sequence of NANOG1 to NANOG2 and NANOGP8. Shade (with solid black) indicates residues that differ from NANOG1. Residues that match the NANOG1 sequence exactly are indicated as '.'. Gaps are indicated as '-'. The last black shaded '.' indicates the STOP codon.

1.6.2. Mouse variants

A total of four variants have been identified in the mouse genome, although they are not related to the human variants. In mouse, *Nanog* variants differ depending on the strain analyzed. Two pseudogenes, denoted *NanogPa* and *NanogPb*, have been found in the C57BL/6 genome (Booth and Holland, 2004). These pseudogenes are critically disrupted by mutations and do not conserve a functional ORF.

Additionally, two retrogenes have been found in different strains, denoted *NanogPc* and *NanogPd*. Both retrogenes conserve a functional ORF when compared to *Nanog*. *NanogPc* presents eight missense mutations and one nonsense mutations which terminates the ORF 49 codons earlier compared to Nanog. *NanogPd* presents six missense mutations and a deletion in

the ORF of 30 bp. NanogPc is found in 129/Ola and *M. spretus*, and NanogPd is found in strains 129/Ola, CBA, in *M. m. castaneus* and *M. spretus*. Further analysis revealed that *NanogPd* conserves its pluripotent function shown by its capacity to confer LIF-independent self-renewal potential to ESCs (Robertson et al., 2006).

2. Expression of pluripotency factors in adult tissues

In adult tissues, a population of stem cells is responsible for homeostasis and repair by replenishing senescent or damaged cells (Blanpain et al., 2007). As other stem cells, these adult or somatic stem cells have the capacity to self-renew and to generate multiple differentiated cell types (Simons and Clevers, 2011). Unlike ESCs, these cells have limited potential and are thought to give rise to cells within the tissue of residence. As shown before, ESCs depend on a cohort of pluripotent genes whose elimination results in commitment into different lineages. Some of these pluripotency genes are also active in rare adult cells, suggesting a common origin of these cell types and showing similarities between embryonic and adult stem cells.

2.1. OCT4

Initial studies of *Oct4* indicated it could have a potential role in adult tissues. Its expression was observed in a variety of adult stem cell compartments, as well as in cultured progenitor cells, where it was thought to play a role in self-renewal analogous to ESCs. This was further supported by analysis of an inducible-*Oct4* mouse model, which showed that ubiquitous overexpression of *Oct4* led to suppression of differentiation in the intestine and dysplasia of epithelia along with expansion of progenitor cells (Hochedlinger et al., 2005).

A recent study using a conditional KO of *Oct4* has put its role in adult tissues in question. In detail, *Oct4* was deleted in specific populations by using different promoters driving Cre expression (used promoter/targeted tissue: *Villin* promoter/intestine, *Mx-1* promoter/bone marrow, *Mx-1* promoter/liver, *Nestin* promoter/brain, liver, and *Keratin15* promoter/hair follicles). Contradicting the previous studies, deletion in all tissues did not lead to any abnormalities or impaired regenerative capacity (Lengner et al., 2007). The previous results were attributed to *Oct4* pseudogenes that have been identified which could be "contaminating" detection of *Oct4* (Liedtke et al., 2007).

2.2. SOX2

Sox2 has been shown to be expressed in multiple adult stem and progenitor cells, with a key role in tissue homeostasis and regeneration (Sarkar and Hochedlinger, 2013). It is expressed in neurogenic regions of the central nervous system, where extracted $Sox2^{+}$ neural progenitor cells

grow in culture and give rise to different cell types, such as neurons, astrocytes and oligodendrocytes, therefore showing multipotency and capacity to self-renew (Ellis et al., 2004). Its importance in neurogenesis in the hippocampus and in neural progenitor cells has been shown by *in vivo* and *in vitro* deletion (Cavallaro et al., 2008; Favaro et al., 2009; Ferri et al., 2004; Suh et al., 2007). $Sox2^{+}$ cells have also been detected in a variety of tissues, associated to its expression pattern during development. It is expressed in stem and progenitors cells of many tissues, including the adult retina (Taranova et al., 2006), trachea (Que et al., 2009), tongue epithelium (Okubo et al., 2009), testes, forestomach, trachea, anus, cervix, esophagus and lens (Arnold et al., 2011; Juuri et al., 2012). In the skin, it is expressed in the dermal papilla of the hair follicle (Biernaskie et al., 2009; Driskell et al., 2009), known to constitute a well-known epidermal stem cell population (Driskell et al., 2011).

Its different roles during embryonic development and adult tissues appear to be regulated by binding to other transcription factors. In ESCs, SOX2 binding to NANOG and OCT4 creates a core transcriptional network in charge of maintaining self-renewal and inhibiting differentiation of these cells (Boyer et al., 2005; Kim et al., 2008). In neurons, where NANOG and OCT4 are not present, SOX2 binds to transcription factor Brn2 to activate *Nestin*, known to be associated to neural progenitor cells (Tanaka et al., 2004). In stratified epithelial cells, SOX2 interacts with epithelial stem cell transcription factor p63, as shown by a recent study where both bind and have overlapping genomic targets in squamous cell carcinoma cells (Watanabe et al., 2014a).

3. Expression of pluripotency factors in cancer

In addition to NANOG, many other reprogramming factors, notably including SOX2, KLF4, OCT4, LIN28 or MYC, are overexpressed in human cancers and, in some cases, play a driving oncogenic role (Suva et al., 2013). It is hypothesized that reprogramming factors in cancer contribute to cellular plasticity and this could be a critical oncogenic feature (Suva et al., 2013).

3.1. NANOG

Expression of NANOG in cancer was first detected in germ cell tumors (Chambers and Smith, 2004; Sperger et al., 2003). The oncogenic potential of NANOG has also been demonstrated in many *in vitro* studies, where it has been shown to promote proliferation, xenograft growth, and migration and invasion **(Table 1)**. Many of these studies have also determined the *NANOG* variant that is being expressed. Several studies indicate NANOGP8 and not NANOG1 could be expressed in cancer **(Table 1)**. Its tumorigenicity has been demonstrated *in vivo* in breast cancer where overexpression of *Nanog* in MMTV-Wnt-1 mice accelerates mammary tumorigenesis and metastasis (Lu et al., 2013).

Table 1. Expression of Nanog in cancer

Origin	Effect	NANOG Targets	Upstream regulators of NANOG	NANOG variant distinction	Reference
Breast and ovarian carcinoma	Chemoresistance	MDR-1	Hyaluronan(HA)-CD44 pathway	No	(Bourguignon et al., 2008)
Ovarian carcinoma	Chemoresistance and CSC population	-	miR-214	No	(Xu et al., 2012)
Breast carcinoma	Cell cycle progression, G0/G1 arrest, and proliferation	Cyclin D1	-	No	(Han et al., 2012)
Ovarian carcinoma	Migration and invasion	Inhibition of FOXJ1 and Ecadherin	-	No	(Siu et al., 2013)
Colorectal carcinoma	Tumor formation and metastasis	-	-	NANOGP8	(Zhang et al., 2012b, 2013)
Colorectal carcinoma	Invasion and metastasis	Snail/Slug	-	No	(Meng et al., 2010)
Colorectal carcinoma	Tumorigenesis	-	c-Jun and ß-Catenin/TCF4	NANOGP8 and NANOG1	(Ibrahim et al., 2012)
Colorectal carcinoma	Tumor initiation		EpCAM	-	(Lin et al., 2012)
Colorectal carcinoma	Immune evasion	Tcl1a		-	(Noh et al., 2012)
Gastric carcinoma	Overexpression in gastric tumors	-	-	NANOGP8 and NANOG1	(Zhang et al., 2010)
Head and neck SCC	Chemoresistance	miR-21	Hyaluronan(HA)-CD44 pathway PKCε	No	(Bourguignon et al., 2012a) (Bourguignon et al., 2009)
Head and neck SCC	Chemoresistance	miR-302	Hyaluronan(HA)-CD44 pathway	No	(Bourguignon et al., 2012b)
Head and neck SCC	CSC with EMT features	-	GSK3ß	No	(Shigeishi et al., 2013)
Hepatocellular carcinoma	Invasion and metastasis	Nodal/Smad3	-	No	(Sun et al., 2013)
Hepatocellular carcinoma	Chemoresistance, self-renewal, and tumorsphere formation	IGF2 and IGFR1	-	No	(Shan et al., 2012)
Lung carcinoma	EMT and CSC properties	ABCB1	-	No	(Chiou et al., 2010)
Prostate carcinoma	Tumorigenesis	-	-	NANOGP8	(Jeter et al., 2009)
Glioblastoma	CSC proliferation, motility and tumor survival	FAK	-	No	(Ho et al., 2012)
Glioblastoma	Self-renewal and tumorigenesis	GLI1	Hedgehog/GLI1	NANOGP8 and NANOG1	(Zbinden et al., 2010) (Po et al., 2010)

Interestingly, *Nanog* has been previously associated with the important tumor suppressor p53 in ESCs. TP53 gene is mutated or deleted in 50% of human tumors (Hollstein et al. 1991). In ESCs, p53 can suppress *Nanog*'s expression either by direct or indirect mechanisms; directly by binding to *Nanog*'s promoter or indirectly by recruitment of a H3 histone deacetylase (Lin et al. 2004). Further studies have also linked this to cancer, where NANOG and p53 establish a functional negative loop in glioblastoma (Zbinden et al., 2010).

Another interest for NANOG in cancer is the search for cancer stem cell markers. Common properties between cancer stem cells (CSC) and embryonic stem cells point out to ESC markers as CSC markers. NANOG and OCT4 are among these. Distinctively, NANOG overexpression is highly correlated with malignancy and chemoresistance of squamous cell carcinomas

(Bourguignon et al., 2012a; Bourguignon et al., 2012b; Chang et al., 2013; Chen et al., 2013; Du et al., 2012; Watanabe et al., 2014b; Ye et al., 2008), where there is a lack of *in vivo* studies.

3.2. OCT4

OCT4 was identified to have a key role in driving germ cell tumors and is currently used as a diagnostic marker for these tumor types (Gidekel et al., 2003).

3.3. SOX2

SOX2 has been demonstrated to be an oncogene in epithelial cancers by multiple studies. Specifically, the SOX2 locus is amplified in human squamous cell carcinomas (SCC) of the lung and esophagus, as well as in small cell lung carcinomas (Bass et al., 2009; Rudin et al., 2012); a clear feature of an oncogene.

SOX2 overexpression alone in progenitor cells of epithelia is insufficient to drive tumorigenesis. A recent study has shown SOX2 requires activated STAT3 to drive its tumorigenic role in SCC. Overexpression of SOX2 in the basal layer of epithelia demonstrated SCC arise only in the forestomach, since expansion of the progenitor cell population leads to exposure to the bile acid-rich environment and posterior inflammation. Inflammation together with SOX2 expression is responsible for carcinoma formation (Liu et al., 2013).

4. Squamous cell carcinoma

Squamous cell carcinomas (SCC) are the most common cancers from ectodermal origin with more than 300,000 deaths per year (Goon et al., 2009; Madan et al., 2010). This type of cancer arises from squamous cells, which serve as a barrier in the skin, esophagus, lung and cervix. These epithelia are constantly being exposed to environmental aggressions (specially the skin and the oral mucosa epithelia), and therefore must be in constant renewal. This renewal relies on a population of adult epithelial stem cells, which need to constantly self-renew to maintain the homeostasis of this tissue. Specifically, progenitors in the basal layer differentiate as they migrate towards the surface ultimately turning into keratinized cells. In the case of the epidermis, different populations of stem cells have been extensively characterized **(Figure 8)** (Arwert et al., 2012; Barker et al., 2010; Blanpain and Fuchs, 2009). One of the common characteristics of squamous neoplasias is the alteration of differentiation, as well as thickening of the epithelium with increased proliferation. Although there are many subtypes of SCC, depending on their site of manifestation and with different risk factors, they are all likely to share disruption of similar pathways that coordinate the squamous differentiation and proliferation pattern (Janes and Watt, 2006).

4.1. Cutaneous squamous cell carcinoma

Non-melanoma skin cancer is the most frequent malignancy in fair-skinned populations. Of these diagnosed non-melanoma skin cancers, 80% were basal cell carcinomas (BCC) and 20% were cutaneous squamous cell carcinomas (cSCC) (Lomas et al., 2012). Unlike almost all BCCs, cSCCs are associated with a substantial risk of metastasis (Alam and Ratner, 2001). Exposure to sunlight is one of best-known causes of this type of cancer, specifically there are studies showing that ultraviolet radiation induces p53 mutations by formation of thymidine dimers in invasive cSCC (Backvall et al., 2005; Brash et al., 1991).

4.2. Esophageal squamous cell carcinoma

Esophageal carcinoma is the sixth leading cause of cancer-associated mortality with an overall 5-year survival of 15%-20%. Squamous cell carcinoma is the most prevalent form of esophageal carcinoma worldwide, although adenocarcinoma also has high incidence in many western countries. The major risk for esophageal squamous cell carcinoma (ESCC) is alcohol and tobacco consumption. One of the main problems is that most patients present metastasis at the time of diagnosis, and treatment consists in combination approaches of surgery and chemotherapy (Pennathur et al., 2013).

In squamous epithelia, proliferation is confined to the basal layer, which gives rise to differentiating epithelial layers (suprabasal layer). Normally, in epithelial tissues a small group of stem cells is responsible for tissue homeostasis **(Figure 8)** (Blanpain and Fuchs, 2009). The esophageal epithelium seems to be unique, since to date, a population of quiescent stem cells that maintains homeostasis and repair has not been identified. Instead, a single layer of progenitor cells equally contributes to both these processes (Doupe et al., 2012).

4.3. Head and neck squamous cell carcinoma

This type of SCC includes lesions in the oral cavity, larynx and pharynx, and about 550,000 new cases are diagnosed each year (Siegel et al., 2012). Although tobacco and alcohol are a risk factor for this disease, currently the major risk factor is infection with oncogenic human papilloma virus (HPV) (D'Souza et al., 2007). This type of cancer arises from the growth of a single stem cell or from a population of tumor initiating cells that have acquired self-renewal capacity (Lobo et al., 2007). The most frequent driving genes that have been identified are *TP53* (mutated in 50-80% of the cases), *CDKN2A*, *PI3KA*, *PTEN*, *HRAS* (Molinolo et al., 2009), and *NOTCH1* (acting as a tumor suppressor in this context) (Agrawal et al., 2011; Stransky et al., 2011).

4.4. DMBA/TPA chemical carcinogenesis as a model for squamous cell carcinoma

The use of multi-stage chemical carcinogenesis models is one of the best models to study *in vivo* the development of tumors, and to evaluate the impact of genetic manipulation on tumor initiation, promotion and progression (Abel et al., 2009).

The initial stage involves provoking mutations in key genes in epidermal progenitor cells by exposing them to a chemical mutagen. 7,12-Dimethylbenz(a)anthracene (DMBA) is a frequently used initiating agent. DMBA causes initiation by mutations that activate *Hras* in most cases (Balmain et al., 1984; Quintanilla et al., 1986), although mutation of *Kras* has also been observed. To give rise to tumors, the cellular targets of these mutations are required to be epidermal stem cells, specifically the hair follicle bulge stem cells (Lapouge et al., 2011; White et al., 2011) and/or the interfollicular epidermis (IFE) stem cells (Brown et al., 1998; Vitale-Cross et al., 2004) **(Figure 8 and 9)**. In humans, mutations in the Ras family are found in 30% of all cancers, and this is also extends to SCCs (Corominas et al., 1991; Spencer et al., 1995; van der Schroeff et al., 1990). There are numerous studies showing that activation of different pathways in the skin stem cells leads to different tumor phenotypes. For example, expression of ßcatenin induced hair-follicle tumors (Gat et al., 1998); expression of GLI1 or GLI2 leds to formation of basal cell carcinomas (Grachtchouk et al., 2000; Nilsson et al., 2000); and in agreement with the above, Hras has specifically shown to induce keratoacanthomas, papillomas and squamous carcinomas (Brown et al., 1998).

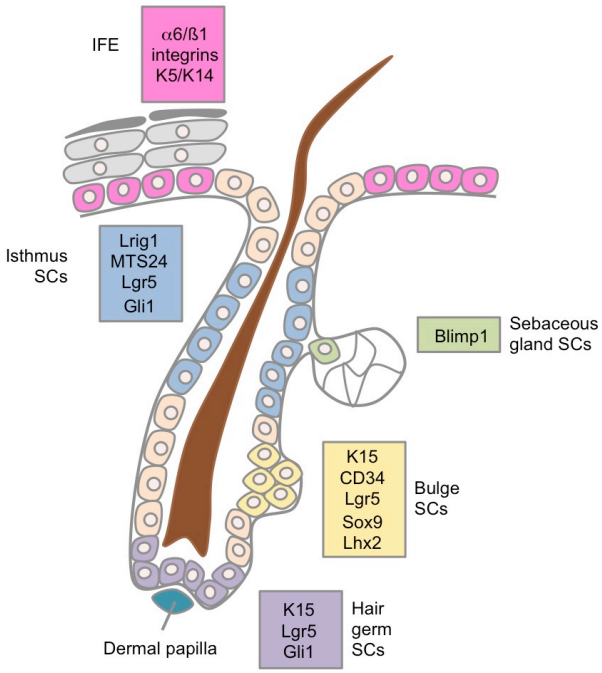


Figure 8. Skin stem cell populations.
Schematic representation of the skin epidermis with the different stem cell (SC) compartments and markers identifying each population. Bulge SCs are multipotent and reside in the non-cycling portion of the hair follicle. The interfollicular epidermis (IFE) also has SCs in the basal layer of the epidermis. There are also sebaceous gland and isthmus progenitor cells between the sebaceous gland and the bulge. Modified from (Alcolea and Jones, 2014).

Expression of activating mutations of *Ras* is not sufficient to induce SCC. The second step that is required in chemical carcinogenesis is the promotion stage. This stage consists on amplifying the population of mutated cells by using chemical agents or wounding that promote not only proliferation, but increased self-renewal of the stem cell population and inflammation (Perez-Losada and Balmain, 2003). One of the best-known tumor promoting agents is the phorbol ester 12-O-tetradecanoylphorbol-13-acetate (TPA). Further evidence that mutations must occur in the stem cell population is shown by the fact that after DMBA treatment, TPA treatment even after one year can still promote papilloma formation (Morris, 2000; Morris et al., 1997; Morris et al., 1986). One of the main targets of phorbol esthers is the protein kinase C (PKC) pathway, particularly of two of its isoenzymes PKCα and PKCε. One of PKC's downstream effects is the activation of Ras and the ERK cascade which promotes proliferation (Griner and Kazanietz, 2007). PKCα and PKCε have also shown to mediate inflammation through overexpression of CXC chemokine receptor 2 (CXCR2) in keratinocytes, and through tumor necrosis alpha (TNFα) and activation of Stat3 respectively. Many mechanisms have been studied for TPA tumor promotion, but in summary it is necessary for skin tumor development by inducing inflammation and proliferation (Rundhaug and Fischer, 2010). After repeated application of the phorbol esther TPA, papillomas arise after some weeks. Papillomas are benign tumors consisting on clonal outgrowths of the skin with a stromal core and hyperplastic epidermis (Yuspa, 1998).

The last step in chemical carcinogenesis assays is progression of papillomas to invasive SCCs or what is known as the conversion stage. This occurs spontaneously and requires additional events such as loss of tumor suppressors or other promoting factors independent of TPA. One example is the loss of TP53, where DMBA/TPA experiments in p53 KO mice show enhanced malignant progression (Kemp et al., 1993) **(Figure 9)**. The conversion rate also varies with the genetic background. In the FVB strain up to 50% of papillomas convert to SCCs, meanwhile in the C57BL/6 strain this is a very rare event (Hennings et al., 1993). Additional analysis has shown that *Mus musculus* inbred mice are much more susceptible than the *Mus spretus* strain. By comparing resistant *M. spretus* to the FVB strain, different networks were identified for susceptibility, including mitosis and cell proliferation genes; and interleukin 1 antagonist activity (Quigley et al., 2009).

SCCs can yet progress further undergoing epithelial-mesenchymal transtion (EMT) and turn into undifferentiated spindle cell carcinomas (SpSCC) (Klein-Szanto, 1989). Elevated *Hras* and TGFß signaling seem to mediate this transition (Cui et al., 1996; Oft et al., 2002). Recently, these SpSCC have been identified to have frequent *Ink4/Arf* deletions and upregulated EMT transcription factors (Wong et al., 2013) **(Figure 9)**.

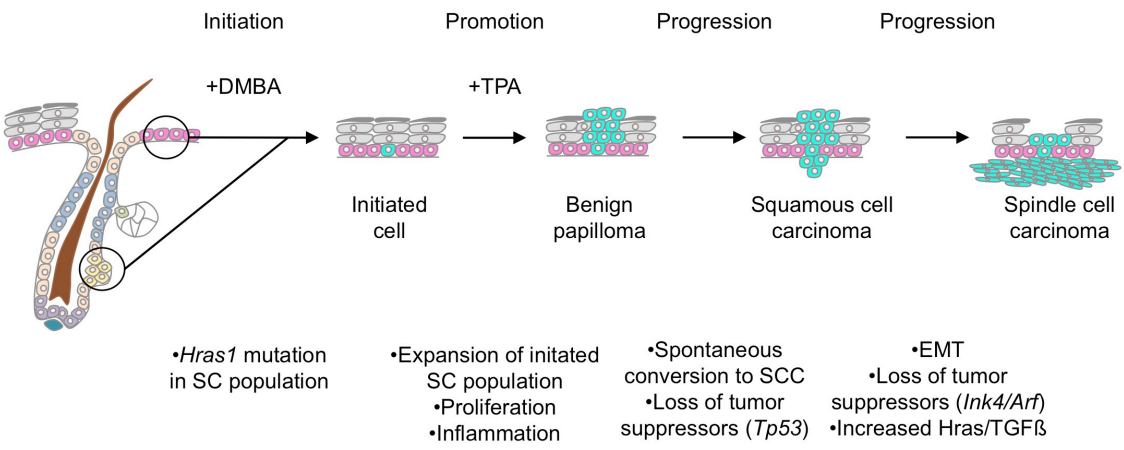


Figure 9. DMBA/TPA chemical carcinogenesis protocol in mice.
Initiation starts with a single topical application of the mutagen DMBA. Repeated topical applications of TPA during 10-40 weeks after initiation promote proliferation and inflammation expanding the initiated stem cell (SC) population. After 6-12 weeks of TPA application, papillomas start appearing. Further mutations and driving events cause a fraction of papillomas to spontaneous convert to SCC. Lastly, EMT and increased TGFß and Hras signaling promote a final progression stage to a very invasive carcinoma known as spindle cell carcinoma.

The importance of the chemical carcinogenesis assay is that it can model human cancer. Human cancer is also a multi-sequential process with different associated events leading to neoplasias. In this model, initiation occurs in a single stem cell, which could also resemble the cancer stem cell model and posterior progression to malignant carcinoma is an additional step in the tumor evolution.

OBJECTIVES

1. Analyze the expression of transcription factor *Nanog* in adult tissues.

2. Determine the role of NANOG in adult tissues and homeostasis using a ubiquitous inducible-*Nanog* transgenic mouse model.

3. Determine the role of NANOG in squamous cell carcinomas by using an epithelia-specific inducible-*Nanog* transgenic mouse model.

4. Determine the role of *NANOG* in epithelial cells and human squamous cell carcinoma cell lines.

5. Analyze the expression of human *NANOG* variants in cancer.

6. Analyze the capacity of human *NANOG* variants to induce reprogramming.

OBJETIVOS

1. Analizar la expresión de NANOG en tejidos adultos de ratón.

2. Determinar el papel de *Nanog* en tejidos adultos utilizando un modelo de ratón modificado genéticamente que sobreexpresa *Nanog* de forma ubicua.

3. Determinar el papel de *Nanog* en carcinogénesis epitelial utilizando un modelo de ratón que sobreexpresa *Nanog* en la capa basal del epitelio.

4. Determinar el papel de NANOG en células epiteliales y derivadas de carcinomas de células escamosas humanas.

5. Determinar la expresión de las variantes humanas de *NANOG* en cáncer.

6. Elucidar la capacidad de las variantes humanas de *NANOG* para inducir reprogramación de células diferenciadas de humano y de ratón.

MATERIALS AND METHODS

1. Animal Experimentation

1.1. Transgenic mouse model

Mice were housed at the specific pathogen-free (SPF) barrier area of the Spanish National Cancer Research Centre (CNIO), Madrid. Mice were observed on a daily basis and sacrificed when they showed overt signs of morbidity in accordance to the *Guidelines for Humane Endpoints for Animals Used in Biomedical Research* from the Council for International Organizations of Medical Sciences (CIOMS). All animal procedures were performed according to protocols approved by the CNIO-ISCIII Ethics Committee for Research and Animal Welfare (CEIyBA). We used three different transgenic mouse models, (1) the *tetO-Nanog* mouse (kindly provided by Konrad Hochedlinger and Rudolf Jaenisch) contains the *Nanog*-cDNA under the control of the doxycycline (DOX)-responsive promoter (*tetO*) inserted downstream of the *Col1a1* locus, as described elsewhere for the generation of OCT4 inducible mice (Hochedlinger et al., 2005); (2) the *Rosa26::rtTA* (kindly provided by Konrad Hochedlinger and Rudolf Jaenisch) mouse contains the reverse tetracycline transactivator (*rtTA*) inserted within the ubiquitously expressed *Rosa26* allele (Hochedlinger et al., 2005); and (3) the *K5-rtTA* mouse (kindly provided by Silvio Gutkind) contains a transgene with the reverse tetracycline transactivator (*rtTA*) cloned downstream from the bovine cytokeratin promoter (*K5*, specifically expressed in the basal layer of stratified epithelia) (Vitale-Cross et al., 2004).

The ubiquitous *Nanog*-inducible mice (Results Part I) correspond to the genotype $Col1a1^{tetO\text{-}Nanog/+}$;$Rosa26^{rtTA/+}$ (abbreviated as TG), and as a control, $Col1a1^{+/+}$;$Rosa26^{rtTA/+}$ or $Col1a1^{tetO\text{-}Nanog/+}$;$Rosa26^{+/+}$ mice were used (abbreviated as CTR). Mice between 3-8 months of both sexes were used for the experiments with a mixed 129Sv/C57BL6 background. Transgene expression was induced by addition of doxycycline to the drinking water of mice, two doses were used, 0.2 mg/ml (LOW DOX) and 2 mg/ml (HIGH DOX) in a 7.5% sucrose solution. For long-term treatment using LOW DOX, a 2% sucrose solution was used. Doxycycline containing water was changed every 2-3 days.

The epithelia specific *Nanog*-inducible mice (Results Part II) correspond to the genotype $Col1a1^{tetO\text{-}Nanog/+}$; $K5\text{-}rtTA^{tg/.}$ (abbreviated as TG), and as a control $Col1a1^{tetO\text{-}Nanog/+}$ (abbreviated as CTR). Mice carrying these alleles were backcrossed at least five time with wild type 100% pure C57BL/6 mice to have them in this background. Littermate mice between 5-8 months of both sexes were used for all experiments, except for the chemical carcinogenesis assay where mice were between 8-10 weeks old at the start of the treatment.

1.2. Topical TPA application

To induce epidermal proliferation, tail skin from a group of five mice per genotype (CTR and TG) in the telogen (resting) phase of the hair cycle was topically treated every 48 hr with TPA (20 nmol in acetone) (cat # P8139; Sigma) for a total of four doses. Five control mice of each genotype were treated with acetone alone. Mice were sacrificed 24 hr after the last TPA treatment, and the tail skin was analyzed.

1.3. Chemical carcinogenesis assay

Twelve age-matched (8-10 weeks old) mice of each genotype, CTR and TG were given doxycycline in their drinking water, which was changed every 2-3 days throughout the entire treatment. 48 hours after, a single dose of 100 nmol of DMBA (cat # D3254; Sigma) in 0.2 ml acetone was topically applied to the shaved dorsal skins to induce DNA mutation. One week later, 17 nmol TPA (cat # P8139; Sigma) in 0.2 ml acetone was administered topically to the same area of the shaved dorsal skin with two administrations per week (Tuesday and Friday) for 15 weeks. Fifteen weeks after this, mice were sacrificed for histopathological analysis. Mice were evaluated weekly for papilloma development. Only tumors that had attained a size of 1 mm were counted. Mice with tumors larger than 1.5 cm were sacrificed considering humane endpoint before the end of the treatment, and analyzed histopathologically.

1.4. Histolopathology, immunohistochemistry and *in situ* hybridization

Tissue samples were fixed in 10% buffered formalin, embedded in paraffin and sectioned at a thickness of 2.5 mm. Consecutive sections were stained with hematoxylin and eosin (H&E) and processed for immunohistochemistry (IHC) performing antigen retrieval in a VENTANA DISCOVERY XT (Roche) with CC1 buffer and incubating them with the following three antibodies against NANOG: Novus Biologicals, NB100-58842 lot# 5-1 (Part I and Part II, Figures 26, 27, and 28), but the provider discontinued the commercialization of this antibody and for this reason we also used Calbiochem, SC1000 (Part I) and Cell Signaling Technology, D2A3, 8822 (Part II, Figures 31 and 33). Other antibodies used were: phospho-Akt1/PKBa (S473) (Epitomis, EP2109Y, 2118-1), AURKA (Becton Dickinson, 610938), CK5 (Covance, PRB-160P), CK6 (Covance, PRB-169P), E-CADHERIN (BD Biosciences, 610182), Ki67 (Master Diagnóstica, 000310QD), LORICRIN (Covance, PRB-145P), OCT4 (Abcam, Ab19857), P63 (Sigma, P3737), SOX2 (Cell Signaling, 3728), phospho-Stat3 (Tyr705) (Cell Signaling Technology, D23A7, 9145) and VIMENTIN (Cell Signaling Technology, D21H3, 5741). Following incubation with the primary antibodies, positive cells were visualized using 3,3-diaminobenzidine tetrahydrochloride plus (DAB+) as a chromogen. Counterstaining was

performed with hematoxylin. Digital slides were acquired with a MIRAX SCAN (Zeiss), images captured with the Pannoramic Viewer Software (3DHISTECH) and image analysed and quantified with the AxioVision software (Zeiss).

For *in situ* hybridization (ISH) of *Nanog*, mouse tissues were fixed overnight in 4% paraformaldehyde, embedded in paraffin, and cut in 7 µm thick sections. *In situ* hybridization was carried out as previously described (Mallo et al., 2000) using a mouse *Nanog* anti-sense probe. This probe corresponds to the complete coding region of the mouse *Nanog* gene (positions 216 to 1133 of RefSeq NM_028016) cloned as a PCR product in the pCR®II-TOPO cloning vector (Invitrogen). The vector was linearized with *Bam*HI and transcribed using T7RNApol. Digital slides were acquired with a MIRAX SCAN (Zeiss) and images were captured with the Pannoramic Viewer Software (3DHISTECH).

ISH of *miR-21* was performed with a *miR-21* Exiqon probe and following the manufacturers' instructions (Exiqon) using the Discovery XT immunohistochemistry system (Ventana Medical Systems).

2. *In vitro* experimentation

2.1. Cell lines

A detail description of cell lines used is provided in Table 2.

Table 2. Cell lines

Cell line	Origin	Phenotype	Culture conditions	Reference
NTERA2	Human testis malignant pluripotent embryonal carcinoma lung metastasis	Human teratocarcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Andrews et al., 1984)
TE1	Human ESCC	Human ESCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Nishihira et al., 1993; Nishihira et al., 1979)
TE2	Human ESCC	Human ESCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Nishihira et al., 1993; Nishihira et al., 1979)
TE5	Human ESCC	Human ESCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Nishihira et al., 1993; Nishihira et al., 1979)
TE12	Human ESCC	Human ESCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Nishihira et al., 1993; Nishihira et al., 1979)
CAL33	Oral cavity carcinoma (tongue), moderately differentiated, 69 year-old male	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Gioanni et al., 1988)

Cal27	Oral cavity carcinoma (tongue), poorly differentiated. 56 year-old male	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Gioanni et al., 1988)
HN6	Base of tongue, well to moderately differentiated. 54 year-old male	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Easty et al., 1981)
SCC2	Oral cavity SCC	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Grenman et al., 1991; Krause et al., 1981)
SCC40	Tongue SCC	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Grenman et al., 1991; Krause et al., 1981)
SCC38	Primary SCC of glottic larynx	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Grenman et al., 1991; Krause et al., 1981)
SCC42B	Larynx-supraglottic SCC	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Grenman et al., 1991; Krause et al., 1981)
SCC29	Alveolus SCC	Human HNSCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Grenman et al., 1991; Krause et al., 1981)
Hela	Cervix carcinoma	Human cervix adenocarcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Gey et al., 1952)
HepG2	Liver hepatocellular carcinoma	Human hepatocellular carcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Knowles et al., 1980)
MCF-7	Invasive breast ductal carcinoma	Human breast adenocarcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Soule et al., 1973)
MDAMB231	Breast carcinoma, cells derived from metastatic site: pleural effusion	Human breast adenocarcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Cailleau et al., 1974)
NCI-H69	Small cell lung cancer	Human small cell lung carcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Gazdar et al., 1980)
NCI-H727	Lung carcinoid	Human well differentiated neuroendocrine cell line	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Gazdar et al., 1985; Gazdar et al., 1988)
A549	Lung carcinoma	Human alveolar adenocarcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Giard et al., 1973)
U2OS	Bone	Human osteosarcoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Ponten and Saksela, 1967)
SAOS	Bone	Human osteosarcoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Ponten and Saksela, 1967)
HCT-116	Colon cancer	Human	DMEM supplemented with	(Brattain et al., 1981)

		colorectal carcinoma	glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	
T-24	Urinary bladder tumor	Human transitional cell carcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Bubenik et al., 1973)
EJ138	Urinary bladder tumor	Human transitional cell carcinoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Lin et al., 1985)
Hs895T	Melanoma lung metastasis	Human melanoma	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	ATCC
22Rv1	Prostatic carcinoma xenograft	Human prostate carcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Sramkoski et al., 1999)
PC3	Protstatic carcinoma bone metastasis	Human prostate carcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Kaighn et al., 1979)
Du145	Protstatic carcinoma brain metastasis	Human prostate carcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Stone et al., 1978)
LNCAP	Protstatic carcinoma lymph node metastasis	Human prostate carcinoma	RPMI supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Horoszewicz et al., 1983)
MCA3D	Mouse primary keratinocytes treated with DMBA and Ca^{+2}	Mouse immortalized keratinocytes	Ham's F12 supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Kulesz-Martin et al., 1983; Quintanilla et al., 1991)
PDV	Mouse primary keratinocytes transformed with DMBA	Mouse transformed keratinocytes	Ham's F12 supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Fusenig et al., 1978; Quintanilla et al., 1991)
PB	Mouse papilloma induced with DMBA/TPA	Mouse papilloma	Ham's F12 supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Yuspa et al., 1986)
MSC11B9	Mouse SCC induced with DMBA/TPA	Mouse SCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Burns et al., 1991)
Pam212	Mouse primary keratinocytes spontaneously tranformed	Mouse transformed keratinocytes	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Yuspa et al., 1980)
CarB	Mouse carcinoma induced with DMBA/TPA	Mouse SpCC	Ham's F12 supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Buchmann et al., 1991)
CarC	Mouse carcinoma induced with DMBA/TPA	Mouse SpCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Buchmann et al., 1991)

MSC11A5	Mouse carcinoma induced with DMBA/TPA	Mouse SpCC	DMEM supplemented with glutamine 2mM; 10% fetal bovine serum and antibiotic-antimycotic solution	(Burns et al., 1991)

Human embryonic stem cells HES2 (NIH code ES02) were grown in DMEM/F12 supplemented with KnockOut™ serum replacement (KSR, 20%, Invitrogen), non-essential aminoacids (Lonza), β-mercaptoethanol (Gibco), mouse embryonic fibroblasts (MEFs)-conditioned medium (5%) and b-FGF (4 ng/ml, R&D Systems #233-FB) (Grigoriadis et al., 2010).

Mouse ESCs were grown under 3% O2 tension in DMEM (high glucose) supplemented with 15% KnockOut™ Serum Replacement (Invitrogen), leukemia inhibitory factor (LIF) 1000 u/ml, 2i (CHIR99021 (1 μM) and PD0325901 (1 μM), Axon Medchem), non-essential amino acids, glutamax and ß-mercaptoethanol. ESCs were routinely grown on feeder cells. For analyses, ESCs were adapted to gelatin-coated plates for at least 3 passages, and this was done in the absence of 2i. *Nanog*$^{-/-}$ (KO) ESCs (kindly provided by Ian Chambers and Austin Smith) carry a deletion of *Nanog* exons 2 and 3 and a GFP expression cassette under the control of the constitutively expressed *CAG* promoter (Chambers et al., 2007).

For the reprogramming assays, we used MEFs containing a doxycycline-inducible cassette expressing the four Yamanaka factors (Abad et al., 2013) or primary human foreskin fibroblasts.

2.2. Murine embryonic fibroblasts (MEFs) isolation and culture

Primary mouse embryonic fibroblasts (MEFs) of wild-type or of the indicated genotypes (TG refers to transgenic, *Rosa26*$^{rtTA/+}$;*Col1a1*$^{tetO\text{-}Nanog/+}$; and CTR refers to control, *Rosa26*$^{rtTA/+}$;*Col1a1*$^{tetO\text{-}Nanog/+}$) were obtained from C57BL/6 mice as described previously (Palmero and Serrano, 2001) and cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco) supplemented with 10% fetal bovine serum (Hyclone) and antibiotic-antimycotic solution (Gibco).

2.3. Keratinocyte isolation and culture

Primary keratinocytes were freshly isolated from wild-type C57BL/6 mice, p53$^{-/-}$ C57BL/6 mice (Jacks et al., 1994) or TG and CTR K5-*Nanog*-inducible C57BL/6 neonates (days 1–3 post-partum). After dispase (STEMCELL Technologies) treatment the epidermis was separated from the dermis, minced and stirred. The derived cell suspension was then filtered through a sterile teflon mesh (Cell Strainer 0.7 μm, Falcon) to remove cornified sheets and keratinocytes were collected by centrifugation (160 g) and seeded on collagen I pre-coated cell culture plates (BD

Biosciences). Keratinocytes were cultured in Cnt-07 (CELLnTEC) medium supplemented with antibiotic-antimycotic solution (Gibco). Transgene activation was done by adding doxycycline to the medium (1 μg/ml) and changing the medium every 2-3 days.

2.4. Reprogramming of murine cells

MEFs were obtained from an inducible reprogrammable mouse generated in our laboratory that contains a polycistronic cassette with transcription factors Oct4, Sox2, Klf4 and cMyc (OSKM) under the control of the tetO element (*tetO-OSKM*), as well as, the reverse tetracycline transactivator in the Rosa26 locus (*Rosa26::rtTA*) (Abad et al., 2013). Passage 1 MEFs grown in DMEM/10%FBS were plated one day prior to infection ($2x10^5$ cells per well in 6 well gelatin coated plates). MEFs were infected with Nanog/NANOG expressing constructs pMXs-mNanog (Addgene plasmid #13354), (Takahashi and Yamanaka, 2006) pMXs-NANOG1^{K82} (Addgene plasmid #18115), (Lowry et al., 2008) site-directed mutated pMXs-NANOG1^{K82} (as described in the main text), pMXs-NANOGP8 (generated from pMXs-NANOG1^{K82} by triple site-directed mutagenesis) or with pMXs-empty (produced by excising *NANOG1* by EcoRI restriction and re-ligation). 48 hr after the first round of infection, medium was changed to iPSC medium (DMEM (high glucose) supplemented with serum replacement (KSR, 15%, Invitrogen), LIF (1,000 U/ml), non-essential amino acids, glutamax and β-mercaptoethanol), with doxycycline (1 μg/ml) to activate expression of OSKM. Three days after doxycycline addition (or when indicated), we supplemented the medium with inhibitors CHIR99021 (1 μM) and PD0325901 (1 μM) (Axon Medchem). Medium was changed every 36 hr, always in the presence of doxycycline and the two inhibitors. Colonies were scored at day 14.

2.5. Reprogramming of human cells

Reprogramming of primary human foreskin fibroblasts was done as previously described. (Takahashi et al., 2007) Retroviral supernatants were produced in HEK-293T cells grown in DMEM supplemented with 10% fetal bovine serum (DMEM/10%FBS) ($5x10^6$ cells per 100 mm diameter dish) transfected with the ectopic packaging plasmid pCL-Ampho (4 μg) and each of the following retroviral constructs (4 μg): pMXs-KLF4 (Addgene plasmid # 17219), pMXs-SOX2 (Addgene plasmid #17218), pMXs-OCT4 (Addgene plasmid #17217)(Takahashi et al., 2007)), and the NANOG expressing constructs described above (see "Reprogramming of murine cells" section). Transfections were performed using Fugene HD transfection reagent (Promega) according to the manufacturer's protocol. Two days later, retroviral supernatants were collected serially during the subsequent 48 hr, at 12 h intervals, each time adding fresh medium to the cells (10 ml). Primary fibroblasts grown in DMEM/10%FBS had been seeded the previous day of infection ($2x10^5$ cells per well in 6-well gelatin coated plates) and received

1 ml of each corresponding retroviral supernatant (three factors, OCT4, SOX2, KLF4, and the fourth factor corresponding to the NANOG constructs or empty control vector). This procedure was repeated every 12 hr for 2 days (total of 4 additions). The day after infection was completed, medium was replaced by DMEM/10%FBS, and kept for three more days. At day 5, cells were trypsinized and reseeded on feeder plates. At day 6, medium was changed to human ES cell medium (see above under "Cell lines" section). 70% of the medium was replaced daily with fresh medium until colonies were visible and stained for alkaline phosphatase at day 28.

2.6. Knockdown assays

2.6.1. NANOG silencing assays using shRNAs

Lentiviruses containing pLKO.1- scramble shRNA (SCR), and pLKO.1-NANOG-shRNA vectors (TRCN000004885, TRCN000004886 and TRCN000004887) obtained from Open Biosystems (Thermo Scientific Open Biosystems) were produced in 293T cells using the lentivirus packaging plasmids pLP1, pLP2 and pLP/VSVG (Invitrogen), by cotransfecting all plasmids using Fugene HD transfection reagent (Promega) according to the manufacturer's protocol. Cells were plated the day prior to infection, and supernatants were collected 48 hr from 293T cells after transfection. Cell lines SCC38, SCC42B, MCF7, NCI-H727 and 22Rv1 were infected concomitantly with all three lentiviruses that target NANOG to obtain a synergic silencing effect. Cells were selected with puromycin to obtain stable shNANOG cell lines.

2.6.2. NANOG silencing assays using siRNAs

For *NANOG* silencing, TE2 human esophageal cancer cells were transfected with a pool of four siRNA duplexes targeting *NANOG* (ON-TARGETplus Human NANOG (79923) siRNA, SMARTpool, Dharmacon) or with a pool of scramble siRNA duplexes (ON-TARGET*plus* Non-targeting Control Pool, SMARTpool, Dharmacon) using Lipofectamine RNAiMAX in Opti-MEM I (Invitrogen) according to the manufacturer's instructions. Cells were processed for RNA extraction 48 hr after transfection.

2.7. *NANOG* overexpression assays

For human NANOG overexpression, HaCat cells were infected with lentiviruses expressing NANOG1 pSin-EF2-Nanog-Pur (Addgene plasmid #16578) (Yu et al., 2007) and with control empty vector pSin-EF2-empty-Pur (produced by excising *NANOG1* by BamHI restriction and re-ligation). Lentiviruses were produced in 293T cells using the lentivirus packaging plasmids pLP1, pLP2 and pLP/VSVG (Invitrogen), by cotransfecting all plasmids using Fugene HD transfection reagent (Promega) according to the manufacturer's protocol. Cells were plated the day prior to infection, and supernatants were collected 48 hr from 293T cells after transfection.

HaCat cells were infected with the lentiviruses and were selected with puromycin to obtain stable HaCat/NANOG1 or HaCat/EV cells.

For murine NANOG overexpression, murine transformed keratinocyte cell line PB was transfected with the mouse NANOG expressing plasmid pPyCAG-Nanog-IP (Addgene plasmid # 13838) using Fugene HD transfection reagent (Promega) according to the manufacturer's protocol. Cells were processed for RNA extraction 72 hr after transfection.

2.8. Cell proliferation and cell cycle profiling assay

For proliferation and cell cycle assays of HNSCC cell lines, cells were transfected and selected with shRNAs (*SCR* or *NANOG*), or treated for 48 hr with MLN8237 (1 µmol/l) or vehicle (DMSO) and pulsed for 1h with 10µM of EdU (5-ethynyl-2'-deoxyuridine). After EdU pulse, cells were collected, fixed in 4% PFA and processed following the Click-iT protocol (Invitrogen), we used EdU Alexa Fluor-488 and Hoechst 33342 (1 µg/ml) to analyze cell cycle phase. Samples were run in a BD FACSCanto II (Becton Dickinson, San Jose CA) equipped with a 488 nm and 407 nm lines, we used pulse processing to collect at least 10,000 single events. All data was analyzed using FlowJo v9.7.5 (Treestar, Oregon).

A crystal violet assay was used to assess cell proliferation of HaCat infected cells. HaCat/EV and HaCat/NANOG1 cells were seeded in quadruplicate in 96-well cell culture plates. At each time point, the cells were washed with PBS, fixed in 1% glutaraldehyde, and stained with 0.5% crystal violet (Sigma) dissolved in methanol. The dye that stained the cells on the plates was eluted with 10% acetic acid and directly measured with the use of a spectrophotometer microplate reader (BioRad) at a wavelength of 590 nm. Cell number was expressed as the optical density at 590 nm.

2.9. Migration assays

For the *in vitro* migration assay, starved cells were pretreated with 10µg/ml mitomycin C for 2hr to block proliferation and a cell-free area was created by scratching the monolayer with a 200µl pipette tip. Cell migration into the wound area was monitored in serum-free medium or in the presence of chemoattractants 10% FBS. Photographs were taken using a phase-contrast microscope (DIAPHOT 300; Nikon). Keratinocyte migration into the denuded area was monitored for up to 24hr. Quantification of migration was done using using ImageJ software.

2.10. Immunofluorescence

Cells were plated in chambers and fixed in 4% paraformaldehyde for 10 min. Sections were blocked with 1% FBS for 1 hr at r.t., and then incubated for 2 hr at r.t. with N-terminal NANOG (Cell Signaling XP) and C-terminal NANOG (R&D AF1997) in Dako Antibody Diluent with Background Reducing Components. After washing (3 times with PBS 0.1% Triton X-100), slides were then stained with donkey anti-rabbit AlexaFluor555-conjugated secondary antibody for Cterminal NANOG and with goat anti-mouse AlexaFluor488-conjugated secondary antibody for Nterminal NANOG (both secondary antibodies from Invitrogen) in Dako REAL™ buffer, 1 hr at r.t., followed by 4',6-Diamidino-2-phenylindole (DAPI) staining to visualize the nuclei. For NTERA2 cells, additional staining was performed with HCS CellMask™ Deep Red stain (Invitrogen, H32721) to view cell morphology. Slides were mounted with Vectashield antifade medium (Vector Laboratories) before confocal analysis. Confocal microscopy was performed with a TCS SP5 laser scanning spectral microscope (Leica Microsystems) equipped with a Plan-Apochromat 40x/1.2 NA oil objective. 8-bits images were acquired using the Leica LAS AF v.2.1 software (Leica Microsystems). The pictures show the maximum projection of Z-stacks.

2.11. Reprogramming efficiency

For quantification of iPSC generation efficiency the total number of iPSC colonies was counted after staining plates for alkaline phosphatase activity (AB0300 Sigma Alkaline Phosphatase Blue Membrane Substrate Solution and SCR004 AP Detection Kit Chemicon) following the manufacturer's instructions.

3. Biochemical assays

3.1. Multi-*Nanog* PCR and cloning

Total RNA was obtained using RNeasy Mini Kit (Qiagen) following the manufacturer´s instructions. Samples were treated with DNase I (Qiagen) before reverse transcription. Total RNA (5 µg) preparations were treated with "Ready-to-go you-prime first-strand beads" (GE Healthcare) to generate cDNA according to the manufacturer's protocol. Control reactions without reverse transcriptase (-RT) were carried out to exclude DNA contaminations. PCR primers were as follows:

multi-*NANOG* PCR:

multiNANOG-F1, 5'-GACGAGGTCCTGGTCAAG-3'

multiNANOG-F2, 5'- ATCCTGAATCTTAGCTACAAACAG -3'

multiNANOG-R, 5'- GTTGCTCTGCATTGGAAGG -3'

These primers were selected to amplify all *NANOG* paralogs except *NANOGP6* and *NANOGP7*, which were considered too divergent from the rest (see Figure 43). The primers were chosen based on their Tm values relative to all the *NANOG* paralogs under consideration, and their GC content.

Two separate multi-*NANOG* PCR reactions were performed for each cDNA sample: primers F1 and R; or primers F2 and R. The resulting PCR products were gel purified and then ligated into pGEM-T vector (Promega), transformed into DHα5 *E. coli*, and individual clones were sequenced.

3.2. RNA extraction, cDNA synthesis and qRT-PCR

In the case of cultured cells and tissues, total RNA was obtained using RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. In the case miRNA analysis, total RNA was obtained from tissues using the miRNeasy Mini Kit (Qiagen) following the manufacturer's instructions. To generate cDNA, total RNA was reverse transcribed using iScript Advanced First Strand cDNA synthesis kit (BioRad), according to the manufacturer's protocols. Quantitative real-time PCR (qRT-PCR) was performed using Master SYBR Green I mix (Invitrogen) or *GoTaq qPCR* Master Mix (Promega) in an ABI PRISM 7700 thermocycler (Applied Biosystems). Calculation for the values was according to the ΔΔCt method (Yuan et al., 2006) and using *Gapdh/GAPDH* as control.

TaqMan® MicroRNA Assays (Applied Biosystems) were used to quantify miRNAs in samples according to the manufacturer instructions with the TaqMan® Universal PCR Master Mix reagent kit (Applied Biosystems). Normalization was performed using U6 as housekeeping small RNA. qRT-PCR for miRNA was performed in an ABI 7500fast Real-Time PCR System.

Primer sets for quantitative real-time PCR

PCR primers for mouse transcripts were as follows:

Primer		Sequence 5' → 3'
mNanog total	fw	CAAGGGTCTGCTACTGAGATGCTCTG
	rev	TTTTGTTTGGGACTGGTAGAAGAATCAG
mZeb1	fw	GCTGGCAAGACAACGTGAAAG
	rev	GCCTCAGGATAAATGACGGC
mZeb2	fw	CAGGCTCGGAGACAGATGAAG
	rev	CTTGCAGAATCTCGCCACTG
mVimentin	fw	CCACTTTCCGTTCAAGGTCAAG

	rev	CGGCTGCGAGAGAAATTGC
mN-cadherin	fw	CTGATAGCCCGGTTTCACTTG
	rev	CAGGCTTTGATCCCTCTGGA
mTwist1	fw	ACGCTGCCCTCGGACAA
	rev	CCTGGCCGCCAGTTTG
endogenous mNanog	fw	CTTCCCTCGCCATCACACT
	rev	ACAGTCCGCATCTTCTGCTTC
mGapdh	fw	TTCACCACCATGGAGAAGGC
	rev	CCCTTTTGGCTCCACCCT
mPrrx1	fw	TTTGAGCGGACACATTACCC
	rev	GCTCATTCCTGCGGAACTT
mPdgfra	fw	GCCAGACATTGACCCTGTTC
	rev	TCTCGATGGCACTCTCTTCC
mKrt18	fw	GCACTCAGGAGAAGGAGCAG
	rev	CAACAGGCTCCACTTGGTCT
mWnt5a	fw	TGTCTTTGGCAGGGTGATG
	rev	AGCCACAGGTAGACAGCTCG
mVcan	fw	AACTGCTTTCCTGATTGGCA
	rev	GATAACAGGTGCCTCCGTTG
mLrg1	fw	CCATGTCAGTGTGCAGATTC
	rev	AAGAGTGAGAGGTGGAAGAG
mPlet1	fw	CTTGACATCCCAAAGCCAGT
	rev	GACTTTGAGGCTGTGCGATT
mLoxl1	fw	ATGTGCAGCCTGGGAACTAC
	rev	TCACCACGTTGTTGGTGAAG
mH19	fw	GAACAGAAGCATTCTAGGCTGG
	rev	TTCTAAGTGAATTACGGTGGGTG
mSrc2	fw	ACATCCTGGTGGGGGAATAC
	rev	CTGTCCACTTGATGGGGAAC
mPecam1	fw	CGTGTGGACCAACCTTACG
	rev	AAGCCATTCTTGCACAGCTC
mCxcl12	fw	TGTGGCTGCTCATCCTCTTTA
	rev	ACAACCTGGTCCAACTCAGC
mE-cadherin	fw	AGCTTGTGGAGCTTTAGATGC
	rev	TTTTCGGAAGACTCCCGATTCA
mCdc20	fw	TTGGCCATGGTTGGATACTT
	rev	GAGTGCTGTGGATGTGCATT
mAurka	fw	TTCAACTCTCCGTTTGAGCC
	rev	CCAAGTTTGACGAGCAGAGA
mAurkb	fw	TTCCGTAGGACTCTCTGGGA
	rev	TCGCTGTTGTTTCCCTCTCT
mKif20a	fw	TCTTCCTGTCGGTCCAACTC
	rev	TGTCCGAAAGGACCTGTTGT

mBub1b	fw	CCACACATCCAGAGGGTCAT
	rev	TCATGTCCACACTTCAGGGA

PCR primers for human transcripts were as follows:

Primer		Sequence 5' → 3'
hEndogenous NANOG1	fw	TTCATTATAAATCTAGAGACTCCAGGA
	rev	CTTTGGGACTGGTGGAAGAATC
hNANOG1/P8	fw	CACTTCTGCAGAGAAGAGTGTCG
	rev	AGCTGGGTGGAAGAGAACACAG
hTWIST1	fw	CCGGAGACCTAGATGTCATTG
	rev	CACGCCCTGTTTCTTTGAAT
hZEB1	fw	GATGATGAATGCGAGTCAGATGC
	rev	ACAGCAGTGTCTTGTTGTTGT
hZEB2	fw	CAAGAGGCGCAAACAAGCC
	rev	GGTTGGCAATACCGTCATCC
hSNAIL	fw	CGAGCTGCAGGACTCTAAT
	rev	CCACTGTCCTCATCTGACA
hCXCL12	fw	GAGCCAACGTCAAGCATCTC
	rev	TCCACTTTAGCTTCGGGTCA
hLGR5	fw	AACATCAGTCAGCTGCTCCC
	rev	TGTAAAGGCCAGTGAATGCTC
hSOX2	fw	GGGAAATGGGAGGGGTGCAAAAGAGG
	rev	TTGCGTGAGTGTGGATGGGATTGGTG
hWNT5A	fw	AGGTGGAACTGCAGCACTGT
	rev	CTCATGGCGTTCACCACC
hLOXL1	fw	TATGTGCAGAGAGCCCACCT
	rev	GTAGCACCCGCACATCGTAG
hN-CADHERIN	fw	GAGGAGTCAGTGAAGGAGTCA
	rev	GGCAAGTTGATTGGAGGGATG
hΔNp63	fw	CAGACTCAATTTAGTGAGCC
	rev	CTGCTGGTCCATGCTGTT
hGAPDH	fw	AGCCACATCGCTCAGACACC
	rev	GTACTCAGCGGCCAGCATGG
hAURKA	fw	GAGGTCCAAAACGTGTTCTCG
	rev	ACAGGATGAGGTACACTGGTTG
hAURKB	fw	CAGTGGGACACCCGACATC
	rev	GTACACGTTTCCAAACTTGCC
retroviralNANOG1/P8	fw	ATGCAACCTGAAGACGTGTGA
	rev	CCCTTTTTCTGGAGACTAAATAAA

3.3. Site-directed mutagenesis

The plasmid pMXs-NANOG was obtained from Addgene (Addgene plasmid #18115). (Lowry et al., 2008) To obtain the different NANOG variants, site-directed mutagenesis was performed using the *QuikChange*™ Site-Directed Mutagenesis Kit (Agilent) following protocols provided by the manufacturer. Mutations were confirmed by Sanger sequencing of the entire *NANOG1* open reading frame.

3.4. Protein extraction and immunoblots

Nuclear protein extracts were prepared using the Nuclear/Cytosol Fractionation kit (K266-100, BioVision) following protocols provided by the manufacturer. Total lysates were prepared using lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 4 mM CaCl, 1.5% Triton X-100, protease inhibitors and micrococcal nuclease). In the case of tissue extracts, lysates were homogenized using a Precellys 24 tissue homogenizer. For mouse NANOG detection, we used two different antibodies: Novus Biologicals, NB100-58842 lot# 5-1 or Calbiochem, SC1000. For human NANOG detection, we used Cell Signaling, D73G4XP (N-terminal detecting) or R&D Systems, AF1997 (C-terminal detecting). Other antibodies used were: γ-tubulin (Sigma, GTU-88), FLAG (FLAG M2, F1804, Sigma Aldrich), GAPDH (Sigma, G8795), phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (Cell Signaling, 9101), p53 (Leica, NCL-P53-CM5P), histone H3 (Abcam, Ab1791), AURKA (IAK1, BD Biosciences, 610938), p63 (4A4, Novus Biologicals, NB100-691) and SMC1 (Bethyl, A300-055A).

3.5. Gene expression profiling

3.5.1. Microarray

Total RNA was extracted from esophagus using RNeasy Fibrous Tissue Mini Kit (Qiagen), and from liver by combining an initial lysis with Trizol (Invitrogen) with a clean-up with RNeasy Kit (Qiagen). Total RNA was labelled with the One-Color Microarray-Based Gene Expression Analysis (Quick Amp Labeling) kit (Agilent). Samples were hybridized to Whole Mouse Genome 4x44K microarrays (Agilent) by following the manufacturer's instructions. Microarrays were scanned on a G2565C DNA microarray scanner (Agilent) and images were analyzed with Feature Extraction Software (version 10.1.1). The raw data has been deposited in the GEO repository (accession number GSE48370). Microarray background subtraction was carried out using normexp method. To normalize the dataset, we performed loess within arrays normalization and quantiles between arrays normalization. Gene set enrichment analysis (Subramanian et al., 2005) was applied using annotations from Reactome. Genes were ranked

based on limma moderated t statistic. After Kolmogorov-Smirnoff test, only those gene sets showing FDR<0.05 were considered enriched between classes under comparison.

3.5.2. RNA-seq

Total RNA was extracted from papillomas using RNeasy Fibrous Tissue Mini Kit (Qiagen). 1 µg of total RNA, with RIN (RNA integrity number) numbers above 7 (Agilent 2100 Bioanalyzer), were used. PolyA+ fractions were processed using TruSeq Stranded mRNA Sample Preparation Kit (Agilent). The resulting directional cDNA libraries were sequenced for 40 bases in a single-read format (Genome Analyzer IIx, Illumina). The complete set of reads has been deposited in the GEO repository (accession number GSE56566). Reads were aligned to the mouse genome (GRCm38/mm10) with TopHat-2.0.4 (Trapnell et al., 2012) (using Bowtie 0.12.7 (Langmead et al., 2009) and Samtools 0.1.16 (Li et al., 2009); allowing two mismatches and five multihits. Transcripts assembly, estimation of their abundances and differential expression were calculated with Cufflinks 1.3.0 -(Trapnell et al., 2012), using the mouse genome annotation data set GRCm38/mm10 from the UCSC Genome Browser (Karolchik et al., 2014).

GSEAPreranked was used to perform a gene set enrichment analysis of the published EMT and CSC pathways. We used the RNA-seq gene list ranked by statistic, setting 'gene set' as the permutation method and we run it with 1000 permutations. We considered only those gene sets with significant enrichement levels (FDRq-value < 0.25) (Subramanian et al., 2005).

3.6. Chromatin immunoprecipitation

For human ChIP, NTERA2 and TE2 cells were crosslinked with 1% formaldehyde for 15 min at r.t.. Crosslinking was stopped by the addition of glycine to a final concentration of 0.125 M. Fixed cells were lysed in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0 and protease inhibitors) and sonicated. As input, 80 µg of protein extract were reserved. For immunoprecipitation, 800 µg of protein were diluted in dilution buffer (1.1% Triton X-100, 0.01% SDS, 1.2 mM EDTA, 167 mM NaCl and 16.7mM Tris-HCl, pH 8.0 and protease inhibitors), pre-cleared with A/G plus-agarose pre-blocked with BSA (Santa Cruz) and incubated with an antibody against mouse NANOG (R&D Biosystems AF1997) or goat IgG isotype (Jackson ImmunoResearch, 005-000-003). Immune complexes were precipitated with A/G plus-agarose pre-blocked with BSA and washed sequentially with low-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.0), high-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl, pH 8.0), LiCl immune complex wash buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate-Na, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0), and TE buffer

(1 mM EDTA, 10mM Tris-HCl pH 8.0), and finally eluted in elution buffer (1% SDS, 50mM $NaHCO_3$). All samples, including inputs, were subjected to reverse crosslinking, treated with proteinase K and DNase-free RNase, and DNA was extracted with QIAGEN PCR purification column (Qiagen) and resuspended in TE buffer. qPCR was performed in triplicate (technical replicates) in two independent experiments. Negative control corresponds to a genomic region that does not bind NANOG, as a positive control we have used the NANOG promoter, since NANOG is known to bind to its own promoter (Gifford et al., 2013).

For mouse ChIP, wild-type and $Nanog^{-/-}$ ESCs and $p53^{-/-}$ keratinocytes were crosslinked with 1% formaldehyde for 15 min at r.t.. Crosslinking was stopped by the addition of glycine to a final concentration of 0.125 M. Fixed cells were lysed in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0 and protease inhibitors) and sonicated. As input, 60 µg of protein extract were reserved. For immunoprecipitation, 600 µg of protein were diluted in dilution buffer (1.1% Triton X-100, 0.01% SDS, 1.2 mM EDTA, 167 mM NaCl and 16.7mM Tris-HCl, pH 8.0 and protease inhibitors), pre-cleared with A/G plus-agarose pre-blocked with BSA (Santa Cruz) and incubated with an antibody against mouse NANOG (Novus Biologicals, NB100-58842 lot# 5-1) or rabbit IgG isotype (Jackson ImmunoResearch, 011-000-003). Immune complexes were precipitated with A/G plus-agarose pre-blocked with BSA and washed sequentially with low-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.0), high-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl, pH 8.0), LiCl immune complex wash buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate-Na, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0), and TE buffer (1 mM EDTA, 10mM Tris-HCl pH 8.0), and finally eluted in elution buffer (1% SDS, 50mM $NaHCO_3$). All samples, including inputs, were subjected to reverse crosslinking, treated with proteinase K and DNase-free RNase, and DNA was extracted with QIAGEN PCR purification column (Qiagen) and resuspended in TE buffer.

Primer sets for PCR after ChIP:

Primer		Sequence 5' → 3'
Nanog negative control	fw	TTGTGTAGCACCAGGGTCA
	rev	GCCTGGCCTGCAAATTATTA
hNANOG promoter	fw	TCCGGCACAGGAAGTTAGTT
	rev	AGACTCACCACCAGCAGCTT
hPRRX1 promoter	fw	AGACTGACCACTGGATGCTTTTG
	rev	GAAAAGCACATGGGAACCTGTC
hPDGFRa promoter	fw	GGTGCCGTATGCAAATGGTT
	rev	CTTTGTCATGCAGAGTACCTCGATT
hTWIST1 promoter	fw	CTGCACATAAGGTCCGCTCAT

	rev	GGGCGCTTTGCTCAATTA
hZEB2 promoter	fw	GCTGATTACCATGAGAAAGGCTGT
	rev	GAGATTGACAGTCGCGTGGAG
hCXCL12 promoter	fw	TCCCCCAACATTCTTGCTCTT
	rev	GGCTTCAGTGTTTTCAGGATGG
hLGR5 promoter	fw	TGGCTGAGCTCGAATTCTTTTC
	rev	CCATCAGTAGGAATGCACTGTGAG
hPITX2 promoter	fw	CTGGGGGTGAAAAACCACAA
	rev	CCAATCAAATCTGCAGCTCACTT
hmiR-21 promoter	fw	GTTATGCTGAAACTGGGCTGCT
	rev	CAAACCAAGAAATCCCCAGGTAG
mAurka-promoter	fw	CAGAGGGGAGACAAAAGAGGAA
	rev	TTAGCATGAGGCCAAGCAGA
mAurka-intronic	fw	TCTGAGAGTTCAGCATGGGAAG
	rev	TGACTACAAAGGGCAGAGGTACTG
mOct4-promoter	fw	CTCTCGTCCTAGCCCTTCCT
	rev	CCTCCACTCTGTCATGCTCA

3.7. Co-immunoprecipitation

For overexpression studies, U2OS cells were transfected with pSin-EF2-Nanog-Pur (Addgene plasmid #16578) (Yu et al., 2007) and/or deltaNp63alpha-FLAG (Addgene plasmid #26979) (Chatterjee et al., 2008) using the Lipofectamine® LTX Reagent (Invitrogen) according to the manufacturer's protocol. Two days later cells were lysed in lysis buffer (1% Triton, 150 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, pH 7.4 and protease inhibitors) and sonicated. As input, 25 µg of protein extract were reserved. For immunoprecipitation, 500 µg of protein were pre-cleared with protein A/G UltraLink Resin (Thermo Fisher Scientific #53132) and incubated with an antibody against FLAG (FLAG M2, F1804, Sigma Aldrich) or NANOG C-terminal detecting antibody (AF1997, R&D Systems). Immune complexes were precipitated with A/G UltraLink Resin (Thermo Fisher Scientific #53132) and washed sequentially with lysis buffer. Bound immunoreactive proteins were then eluted at 95 °C for 5 min with 20 µl of 1 x Laemmli loading buffer in the presence of 20% β-mercaptoethanol using SigmaPrep columns (Sigma Aldrich). Twenty µl of sample was loaded per lane, and immunoblotting studies were performed as described previously.

For endogenous detection, TE12 cells were grown and processed as above. For immunoprecipitation, 5 mg were pre-cleared with protein A/G UltraLink Resin (Thermo Fisher Scientific #53132) and incubated with an antibody against NANOG N-terminal detecting antibody (Cell Signaling Technology, D73G4XP) or rabbit IgG iaotype (Jackson ImmunoResearch, 011-000-003). Immune complexes and elution was done as mentioned above.

4. Analysis of human cancers

Esophageal and head and neck squamous cell carcinomas were collected and handled anonymously at collaborating institutions (University College London Hospital and Hospital Clinic of Barcelona) with the approval of their institutional review boards and following standard ethical and legal protection guidelines regarding human subjects, including informed consent for all the subjects. The head and neck squamous cell carcinoma tissue array (n=46) was constructed at the Spanish National Cancer Research Centre and included 2 to 4 cylinders for each patient tumour. Consecutive paraffin sections were processed for immunohistochemistry performing antigen retrieval in a Leica Bond III (Leica) and incubating them with antibodies against human NANOG (Cell Signaling, D73G4). Following incubation with the primary antibodies, positive cells were visualized using 3,3-diaminobenzidine tetrahydrochloride plus (DAB+) as a chromogen. Counterstaining was performed with hematoxylin. Digital slides were acquired with a MIRAX SCAN (Zeiss), images captured with the Pannoramic Viewer Software (3DHISTECH). Slides were scored by a trained pathologist. Statistical analyses were performed using the SPSS statistical package (version 18.0).

5. Statistical analysis

Values were expressed as mean ± SD or mean ± s.e.m and differences with P value <0.05 were considered significant ((*) P<0.05, (**) P<0.01, (***) P<0.001). Comparisons between two groups were performed using unpaired two-tailed Student's t-test. For survival curves, we used the Long-rank (Mantel-Cox) test. Time courses were analyzed by two-way analysis of variance (ANOVA). These statistical analyses were performed using GraphPad Prism software.

RESULTS

PART I
Role of NANOG in adult tissues

The results reported in this part I have been obtained in close collaboration with Daniela Piazzolla (CNIO). Other collaborators who have contributed to specific parts are mentioned in the text.

1.1. NANOG expression in stratified epithelia

Our first approach was to analyze if NANOG expression was detectable in adult tissues. We performed the analysis in the C57BL/6 mouse strain, which does not contain *Nanog* pseudogenes that could give rise to protein (Booth and Holland, 2004; Robertson et al., 2006). NANOG immunohistochemical staining was carried out in a wide range of tissues. Our observations show that NANOG is being expressed in the basal layer of stratified epithelia in adult tissues, including the esophagus, forestomach, skin, tongue, vagina, urinary bladder and urothelium, as well as in the myoepithelial layer of the mammary gland and prostate **(Figure 10A)**. Expression was not detected in the liver, lung, skeletal muscle, heart, intestine kidney, testis and brain **(Figure 10B)**. To confirm these results, the group of Miguel Manzanares (CNIC, Madrid) detected the presence of *Nanog* transcripts by *in situ* hybridization with a *Nanog* probe in the epithelial layers of the esophagus, forestomach and skin. As a control, expression was not present in liver **(Figure 10A and 10B)**.

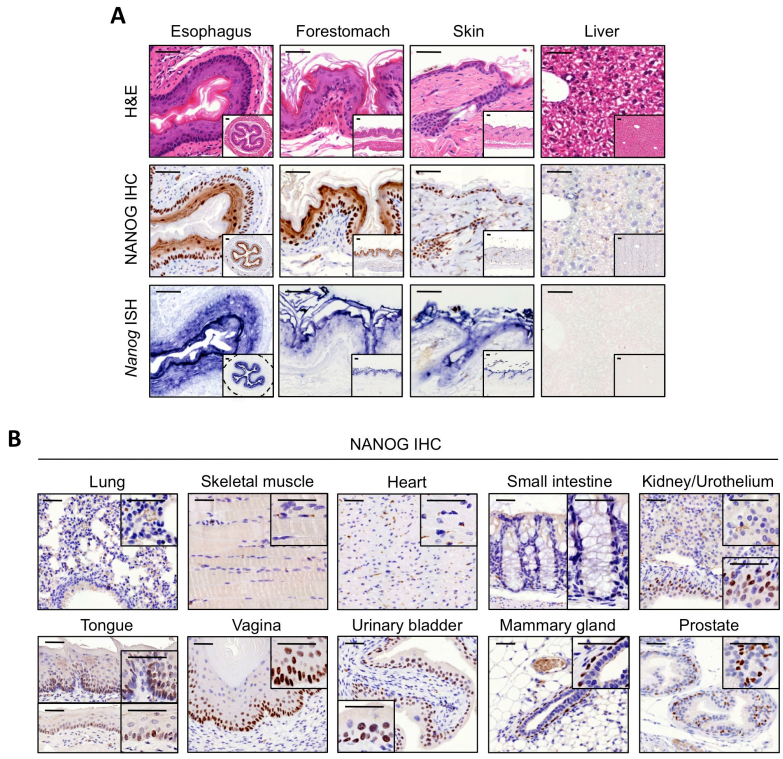


Figure 10. NANOG expression in mouse adult tissues

(A) Hematoxylin and eosin (H&E) staining, immunohistochemistry (IHC) for NANOG, and *in situ* hybridization (ISH) using a *Nanog* RNA probe of the indicated mouse adult tissues.

(B) IHC of NANOG in the indicated adult tissues. Negative tissues: lung, skeletal muscle, heart and small intestine. In the case of kidney, cortex and medulla are negative (upper part of the panel) and the urothelium is positive (lower part of the panel). Positive tissues: tongue (dorsal, upper part; ventral, lower part), vagina, urinary bladder, mammary gland and prostate.

Bars correspond to 50 µm. Two magnifications are shown for each tissue. All samples derive from the C57BL/6 strain.

We decided to compare our results with those of SOX2 and OCT4, which are known partners of NANOG in ESCs (Boyer et al., 2005). As expected, OCT4 was absent in stratified epithelia and liver, and SOX2 was present in esophagus and forestomach, but weak in skin **(Figure 11A)** (Arnold et al., 2011; Lengner et al., 2007). NANOG and SOX2 coincide with the expression pattern of the basal progenitor cell marker P63 **(Figure 11A)**.

Additionally, immunoblots of total lysates from tissues corroborated NANOG expression in esophagus and forestomach, but not in liver or muscle **(Figure 11B)**. To confirm the antibody's specificity, we used wild-type ESCs (WT) and Nanog knock-out ESCs (KO) (Chambers et al., 2007). Of note, NANOG levels in adult tissues are lower than in ESCs **(Figure 11C)**, this is also relative because the basal layer of the epithelia accounts only for a fraction of the tissue. Adult tissues also seem to have different electrophoretic mobility, shown by two distinct bands in the immunoblots. This could be explained by the previously described isoforms or post-translational modifications of NANOG (Saunders et al., 2013).

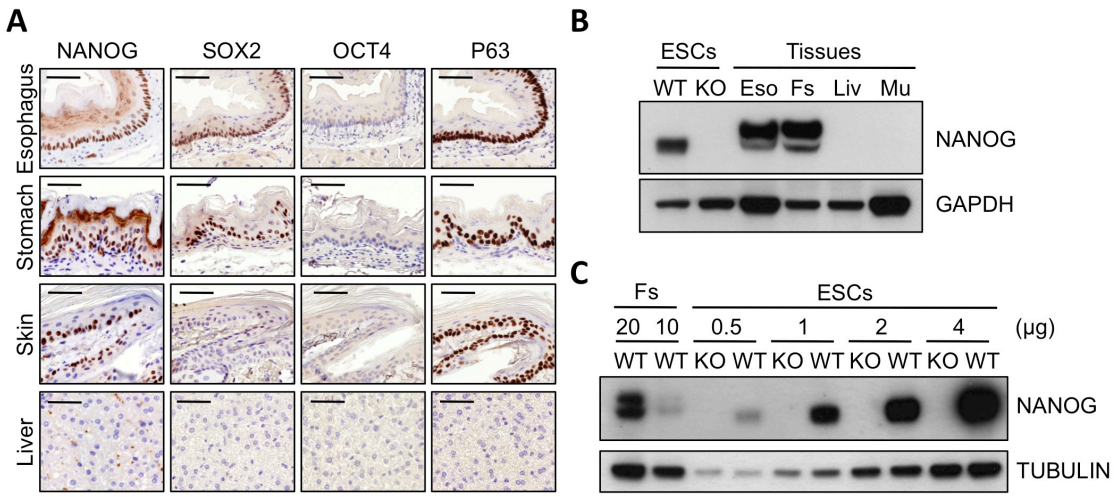


Figure 11. NANOG expression in epithelial tissues

(A) IHC for NANOG, SOX2, OCT4 and P63 of the indicated adult tissues. Stomach corresponds to forestomach.

(B) Immunoblot of NANOG extracts from the indicated tissues. Extracts from wild-type (WT) and Nanog knock-out (KO) ESCs are included as positive and negative controls. GAPDH was used as loading controls. Eso, esophagus; Fs, forestomach; Liv, liver; Mu, muscle.

(C) Immunoblot of NANOG in forestomach (Fs) extracts compared to extracts from WT and Nanog-KO (KO) ESCs. Different amounts of protein were loaded as indicated. γ-TUB is used as a loading control.

Bars correspond to 50 µm. Two magnifications are shown for each tissue. All samples derive from the C57BL/6 strain.

Finally, we wondered about the epigenetic status of the *Nanog* promoter in NANOG-positive epithelia. In particular, the differentially methylated region (DMR) of *Nanog* is hypomethylated in pluripotent cells, such as ESCs, but it is heavily methylated in differentiated cells and in adult tissues, such as brain and testis, that do not express *Nanog* (Blelloch et al., 2006; Hattori et al., 2007; Imamura et al., 2006; Wernig et al., 2007). To investigate the DNA methylation status of the *Nanog* promoter in stratified epithelia, with the help of María Lozano (CNIO) we microdissected the NANOG-positive basal layer of esophagus and forestomach, as well as the

underlying NANOG-negative muscular layers. The extracted DNA was analyzed by the group of Mario F. Fraga (IUOPA, Universidad de Oviedo, Asturias). Briefly, DNA was treated with bisulphite and directly subjected to pyrosequencing. We focused on two central CpG residues of the *Nanog* DMR core region (**Figure 12A**). Importantly, similarly to ESCs, *Nanog* promoter was essentially devoid of methylation in the basal epithelial layers of esophagus and forestomach, whereas their respective adjoining muscular layers were methylated to the same extent as liver or fibroblasts (**Figure 12B**).

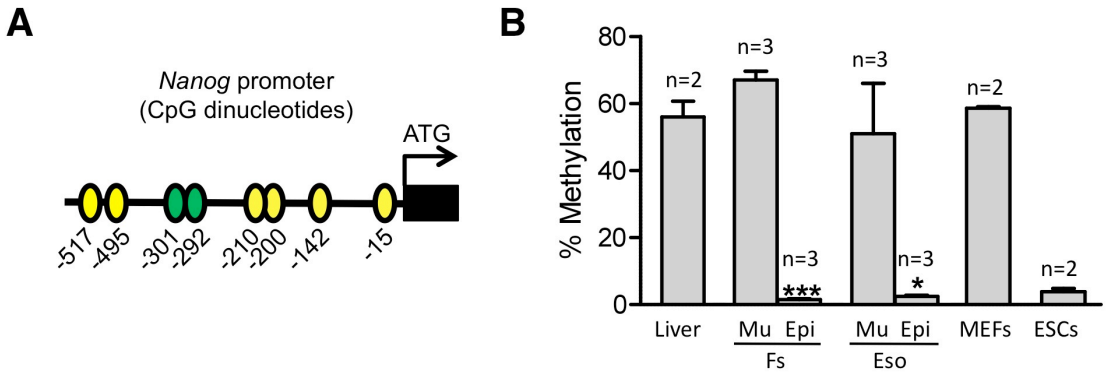


Figure 12. Methylation of the *Nanog* promoter region in adult epithelia

(A) Scheme of the promoter region of the *Nanog* gene. The position of CpG dinucleotides is indicated with ovals. The green ovals indicate the CpG residues analyzed by pyrosequencing (-292 and -301)

(B) Methylation status of the *Nanog* promoter determined by pyrosequencing. Microdissected epithelial (Epi) and muscular (Mu) layers from esophagus and forestomach (n=3 mice) from C57BL/6 mice were analyzed (a total of n=6 samples). Liver was also analyzed as a NANOG-negative tissue. MEFs (2 independent isolates) and ESCs (2 independent lines) were used as positive and negative controls, respectively. Values correspond to mean ± s.e.m. Statistical significance was determined by the two-tailed Student's t test comparing epithelial vs. muscle, or vs. liver (left), or comparing ESCs vs. MEFs (right): ***P<0.001 in all cases.

In summary, multiple independent techniques (immunohistochemistry, *in situ* hybridization, immunoblotting, and DNA methylation) indicate that *Nanog* is expressed in the basal layer of stratified epithelia in adult mice.

1.2. NANOG overexpression causes hyperplasia in stratified epithelia

To interrogate the functional effects of NANOG in adult mice, we used a doxycycline-inducible transgenic mouse model (*Rosa26::rtTA, Col1a1::tetO-Nanog*) kindly provided by Konrad Hochedlinger (Harvard University, Boston, MA, USA) **(Figure 13A)**. These mice are similar to a previously described *Oct4*-inducible mouse model (Hochedlinger et al., 2005). Upon increasing doses of doxycycline (DOX) administration (0.2 mg/ml, LOW; 2 mg/ml, HIGH), we observed increasing levels of NANOG by immunoblotting, qRT-PCR and immunohistochemistry, in all tissues tested except brain and testis that are poorly accessible to doxycycline **(Figure 13B, 13C and 13D)**.

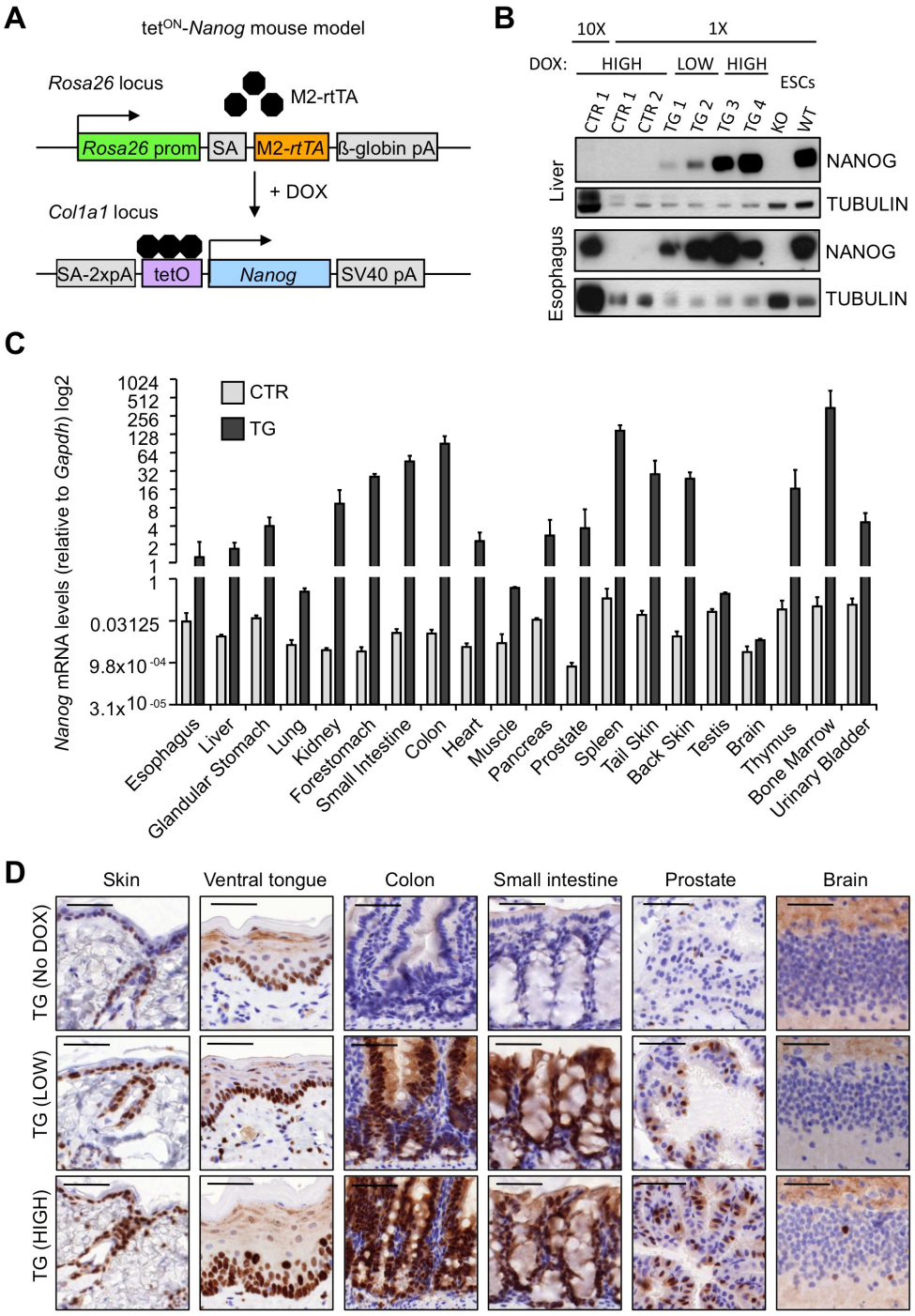


Figure 13. Inducible *Nanog* transgenic mice

(A) Schematic representation of the knock-in alleles generated to produce *Nanog*-inducible mice. The *M2-rtTA* gene was targeted into the ubiquitously expressed *Rosa26* locus. A cassette containing the *Nanog*-cDNA under the control of the doxycycline (DOX)-responsive promoter (*tetO*) was inserted downstream of the *Col1A1* locus. SA, splice acceptor; pA, polyadenylation signal; tetO, tetracycline/doxycycline-responsive operator. Arrows indicate transcriptional start sites.

(B) Immunoblot of NANOG in liver and esophagus of *Nanog*-inducible mice treated with different doses of DOX. Control mice (CTR) correspond to *Rosa26*$^{rtTA/+}$,*Col1a1*$^{+/+}$ and *Nanog*-inducible mice (TG) correspond to *Rosa26*$^{rtTA/+}$,*Col1a1*$^{tetO\text{-}Nanog/+}$. Samples from two independent mice were used for each condition, labelled CTR1, CTR2 and TG1 to TG4, as indicated. Two lanes were loaded with CTR1 extract: 1X, as all the other samples; and 10X, in the leftmost lane. Mice were treated with 0.2 mg/ml (LOW) or 2 mg/ml (HIGH) doxycycline (DOX) for 48 hr. WT and Nanog KO ESCs were used as positive and negative controls. γ-TUBULIN was used as a loading control.

(C) Relative log2 Nanog mRNA levels in the indicated organs determined by qRT-PCR of TG mice (n=3) and CTR mice (n=3) after 48 h after administration of LOW DOX (0.2mg/ml) in the drinking water. Gene expession was normalized to *Gapdh*. (n), number of mice.

(D) Immunohistochemistry (IHC) for NANOG in the indicated tissues (bars indicate 50 µm). TG and CTR mice were treated with 2 dosis of DOX for 96 h. As controls, we used TG mice not treated with DOX (No DOX). Induction of NANOG was evident in all the tissues tested except in testis and brain, likely due to the low accessibility of DOX to these organs.

To assess the long-term impact of NANOG overexpression, we treated adult *Rosa26*$^{rtTA/+}$;*Col1a1*$^{tetO\text{-}Nanog/+}$ mice (abbreviated as TG) and their control littermates carrying only the transactivator *Rosa26*$^{rtTA/+}$;*Col1a1*$^{+/+}$ (abbreviated as CTR) with different doses of DOX and observed their phenotypic changes upon continuous administration. In the case of LOW DOX, half of the TG mice died within 16 weeks, and in the case of HIGH DOX, the median survival was 6 weeks **(Figure 14A)**. Morbidity was associated with a lower body weight **(Figure 14B)**.

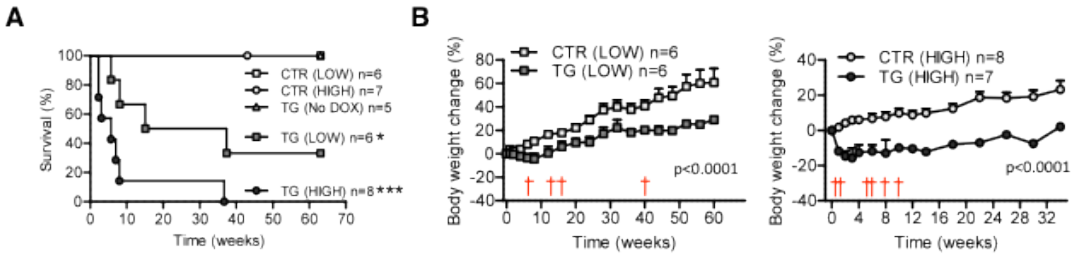


Figure 14. Phenotype of Nanog-inducible mice

(A) Survival curves of *Nanog* inducible mice upon continuous treatment with LOW or HIGH DOX, Control mice (CTR) correspond to *Rosa26*$^{rtTA/+}$;*Col1a1*$^{+/+}$ mice. As an additional control, we include TG mice not treated with DOX (No DOX). Statistical significance was determined by the long-rank (Mantel-Cox) test (n, number of mice). TG (LOW) vs. each of the three controls, **P*<0.05. TG (HIGH) vs. each of the three controls, ****P*<0.001.

(B) Body weight change of CTR and TG mice continuously treated with LOW (left) or HIGH (right) DOX. All values are relative to the body weight at the start of the treatments (time 0). Red crosses indicate sacrificed animals that reached the humane endpoint. Values correspond to mean ± s.e.m. of the indicated number of mice (n). Statistical significance was determined by the two-way ANOVA.

To identify the pathological effects of NANOG overexpression, with the help of Marta Cañamero (CNIO), we performed an extensive histological analysis of TG and CTR mice at their time of death or at the experimental endpoint. The most striking alteration in TG mice was a marked hyperplasia of the esophagus and forestomach **(Figure 15A)**. As previously reported, we also observed thymus atrophy (Tanaka et al., 2007) consequence of Nanog overexpression in hematopoietic stem cells that leads to their erroneous differentiation **(Figure 15B)**. In contrast, no significant changes were observed in major organs, such as liver, lung, heart and kidney. These alterations in the upper digestive tract could impair nutrition and thus explain the reduced body weight.

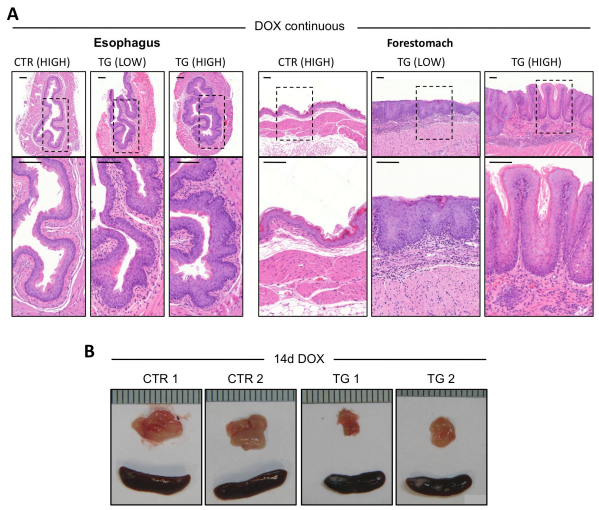


Figure 15. *Nanog* overexpression promotes hyperplasia of stratified epithelia and thymus atrophy

(A) Hematoxylin and eosin (H&E) staining in esophagus (left) and forestomach (right) upon NANOG overexpression showing epithelial hyperplasia. Control *Rosa26*$^{rtTA/+}$;*Col1a1*$^{+/+}$ mice (CTR) and *Nanog*-inducible *Rosa26*$^{rtTA/+}$;*Col1a1*$^{tetO\text{-}Nanog/+}$ mice (TG) were continuously treated with LOW or HIGH DOX. Histological analysis was performed at the humane endpoint of TG mice (approximately 2 months) or at the end of the experiment for CTR mice (11 months). Bars correspond to 100 µm.

(B) Representative images of the spleen and thymus from CTR and TG treated during 14 days with HIGH DOX.

To better define the effects of NANOG overexpression, we analyzed TG and CTR mice that had been treated for 2 weeks with HIGH DOX. We confirmed the induction of hyperplasia in esophagus by NANOG overexpression **(Figure 16)**. Hyperplasia was accompanied by expression of cytokeratin 6 (CK6), which is associated with rapidly proliferating and hyperplastic squamous epithelial tissues (Navarro et al., 1995; Wilson et al., 1990). We also observed an altered differentiation pattern, characterized by an expansion of basal layer markers P63 and CK5 into suprabasal layers, and by a reduced expression of the differentiation marker LORICRIN **(Figure 16)**. Additionally, the ventral side of the tongue was hyperplastic, and the skin presented hyperplastic foci. Other minor alterations were detected in the bladder and intestine, whereas no alterations were found in liver, lung, heart, smooth and skeletal muscle, kidney, prostate and mammary gland.

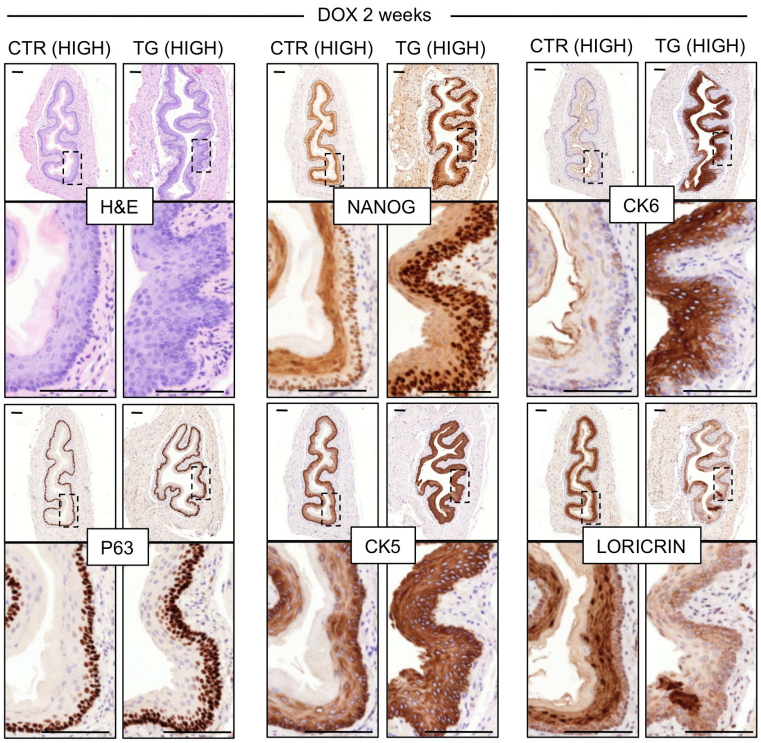


Figure 16. Nanog-associated hyperplasia in the esophagus was accompanied by impaired differentiation
Hematoxylin and eosin (H&E) and immunohistochemistry (IHC) of the indicated proteins in esophageal serial sections of CTR and TG mice treated for 2 weeks with HIGH doxycycline as in Figure 15A. Bars correspond to 100 µm.

The lack of effect of NANOG on the functionality of liver and kidney was confirmed by normal blood levels of alanine transaminase (ALT) and total bilirubin (TBIL) **(Figure 17A and 17B)**.

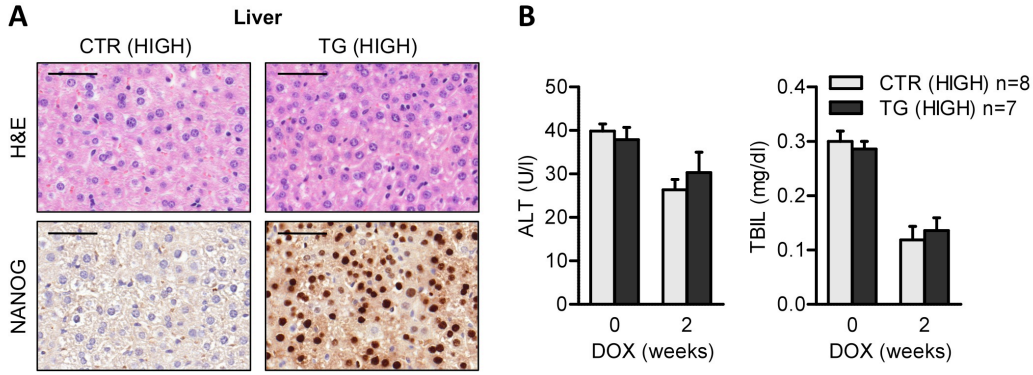


Figure 17. Nanog overexpression did not produce alterations in the liver or kidney function

(A) Hematoxylin and eosin (H&E) and NANOG immunohistochemistry (IHC) in the liver of CTR and TG mice treated with 2 weeks HIGH DOX (bars correspond to 50 µm).

(B) Blood indicators of liver and kidney function in CTR and TG mice treated for 2 weeks with HIGH DOX, as in (A). The indicators measured are alanine transaminase (ALT) and total bilirubin (TBIL). Values correspond to mean ± s.e.m. of the indicated number of mice (n). No significant differences were observed using the two-tailed Student's t test.

In summary, the most noticeable effects of NANOG are restricted to stratified epithelia and, interestingly, these tissues correspond to those with high endogenous expression of NANOG.

1.3. Mitotic activity of NANOG in stratified epithelia

To understand the mechanisms underlying NANOG-induced hyperplasia in stratified epithelia, we treated TG mice and their control littermates with LOW or HIGH DOX for only 48 h to capture early events triggered by NANOG. At this early time point, TG mice presented an increased proliferation, measured by Ki67, in the basal and suprabasal layers of the esophagus and forestomach **(Figure 18)**. In the esophagus, we also observed an increase in DNA damage, measured by phosphorylated H2AX (γH2AX), probably reflecting the induction of replicative stress by NANOG **(Figure 18)**. Interestingly, we could not observe any increase in proliferation or DNA damage in the esophageal muscular layers.

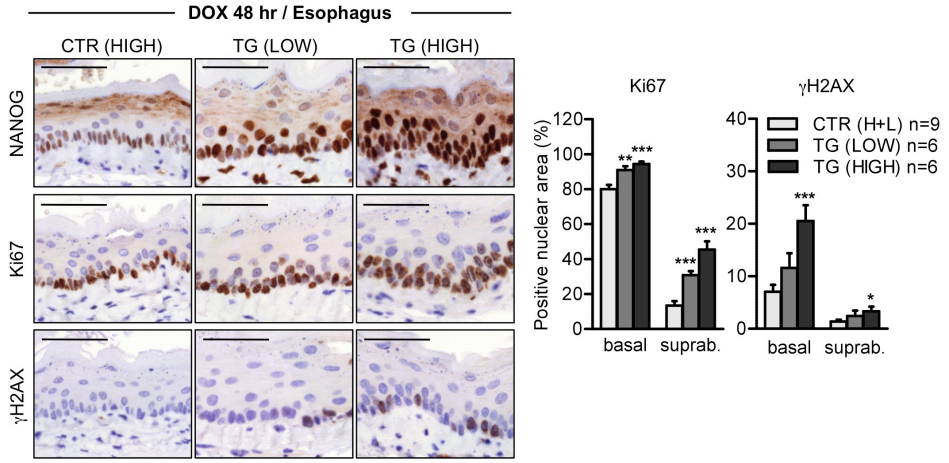


Figure 18. *Nanog* overexpression promotes replicative stress

Immunohistochemistry of NANOG, proliferation marker Ki67, and DNA damage marker γH2AX in esophageal sections of control *Rosa26*$^{rtTA/+}$;*Col1a1*$^{+/+}$ mice (CTR) and *Nanog*-inducible *Rosa26*$^{rtTA/+}$;*Col1a1*$^{tetO\text{-}Nanog/+}$ mice (TG) treated with LOW or HIGH DOX for 48 hr. Bars correspond to 50 µm. The graphs below correspond to the percentage of Ki67 and γH2AX positive nuclear area for each condition in the basal and suprabasal (suprab.) epithelial layers.

Values correspond to mean ± s.e.m. of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: **P*<0.05; ***P*<0.01, ****P*<0.005.

Based on the above observations at such an early time point, we surmised that NANOG is directly implementing a proliferative transcriptional program. To investigate this, we performed gene expression microarray analyses in collaboration with Orlando Dominguez (CNIO) and Gonzalo Gómez López (CNIO), at 48 hr after LOW DOX, in total esophagus and liver, which represent responsive and non-responsive tissues, respectively. Gene set enrichment analysis (GSEA) using public annotations from Reactome revealed that the overwhelming majority of pathways significantly upregulated by NANOG in the esophagus were related to mitosis **(Figure 19A)**, being "M phase" set top upregulated pathway **(Figure 19A and 19B)**. In contrast to the changes observed in esophagus, no pathways were significantly altered in liver. Inspection of the heatmaps for individual genes present in the Reactome "M phase" pathway suggested that NANOG could regulate the expression of pivotal mitotic genes specifically in esophagus but not in liver **(Figure 19C)**. Validation by quantitative PCR was obtained for *Aurka*, *Aurkb*, *Bub1b, Cdc20* and *Kif20a*, which showed upregulation upon NANOG induction in esophagus, but not in liver **(Figure 19D)**. These results indicate that NANOG enforces a proliferative program selectively in the esophagus.

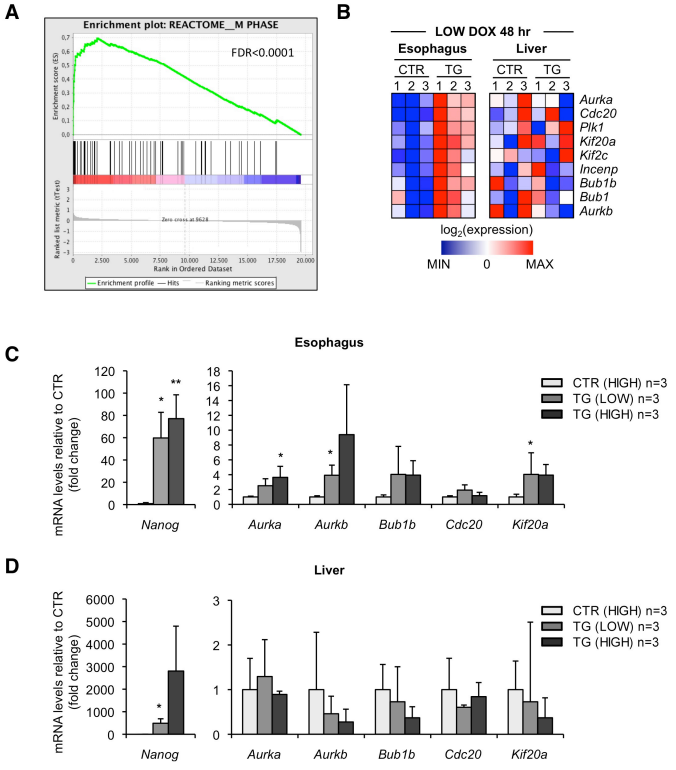


Figure 19. *Nanog* overexpression promotes a mitotic transcriptional program in the esophagus

(A) Gene set enrichment analysis (GSEA) plot of the "M phase" (mitotic phase) REACTOME pathway. GSEA ranks genes according to the ratio between TG and CTR esophagus, both treated with LOW DOX for 48 hr. Red-blue bar: red upregulated in TG and blue downregulated in TG. The false discovery rate (FDR) is indicated.

(B) Heat map of the expression of selected mitotic genes in TG and CTR, as described in A, in esophagus and liver. Values correspond to tissues from three independent mice, labelled 1 to 3, for TG and CTR. The color intensity is proportional to the signal $\log_2$ intensity: red upregulated and blue downregulated.

(C) Relative gene expression in esophagus of *Nanog* and indicated mitotic genes determined by qRT-PCR.

(D) Relative gene expression in liver of *Nanog* and indicated mitotic genes determined by qRT-PCR

In C and D, mice (n=3) were treated as in A. mRNA levels were normalized to *Gapdh* and then compared to CTR levels (fold change). Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: **P*<0.05; ***P*<0.01; ****P*<0.001.

1.4. Aurka is a direct target of NANOG

The above data on the transcriptional effects of NANOG in esophagus were based on the analysis of total esophageal tissue, which includes epithelial and muscular layers. By immunohistochemistry, we observed that NANOG induced an increase in Aurora kinase A (AURKA) positive cells in the esophageal epithelium **(Figure 20A)**, but not in the underlying

muscular layer or in the liver **(Figure 20A and 20B)**. These data indicate that NANOG increases AURKA protein specifically in the esophageal epithelium.

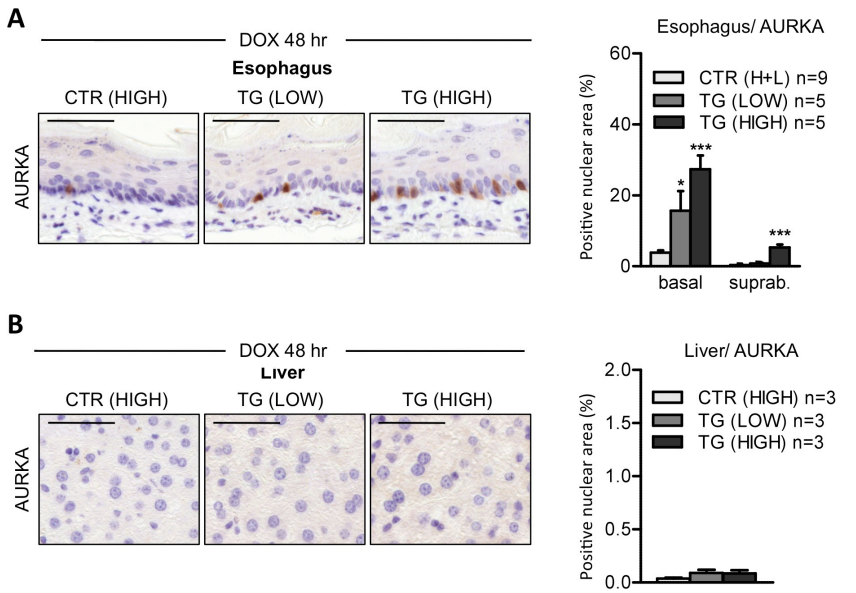


Figure 20. Nanog overexpression induces Aurka upregulation

Immunohistochemistry of AURKA in esophagus (A) and liver (B) of CTR and TG mice. The graphs to the right correspond to the percentage of AURKA positive nuclear area for each condition in the basal and suprabasal (suprab.) epithelial layers of the esophagus, or in the liver. CTR mice treated with LOW or HIGH DOX were pooled (H+L). Values correspond to mean ± s.e.m. of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: **P*<0.05; ****P*<0.001.

To further explore the link between NANOG and aurora kinases, we recapitulated the above *in vivo* observations in a mouse papilloma-derived cell line, PB (Bornachea et al., 2012; Yuspa et al., 1986). Indeed, overexpression of NANOG resulted in a significant transcriptional upregulation of *Aurka* and *Aurkb* mRNA levels (**Figure 21A**). We also wanted to confirm this data in human samples. We used ESCC cells, which have previously been reported to express NANOG (Du et al., 2012; Shimada et al., 2012). Interestingly, transfection of siRNAs in ESCC TE2 cells (Okano et al., 2000) against *NANOG* produced a significant decrease of *NANOG* mRNA (detected with PCR primers specific for *NANOG1* paralog (see Part III, Figure 45 and 47) and this was accompanied by a similar decrease in *AURKA* and *AURKB* mRNA levels **(Figure 21B)**. Similar data were obtained in HNSCC cells (see Part II, Figure 27).

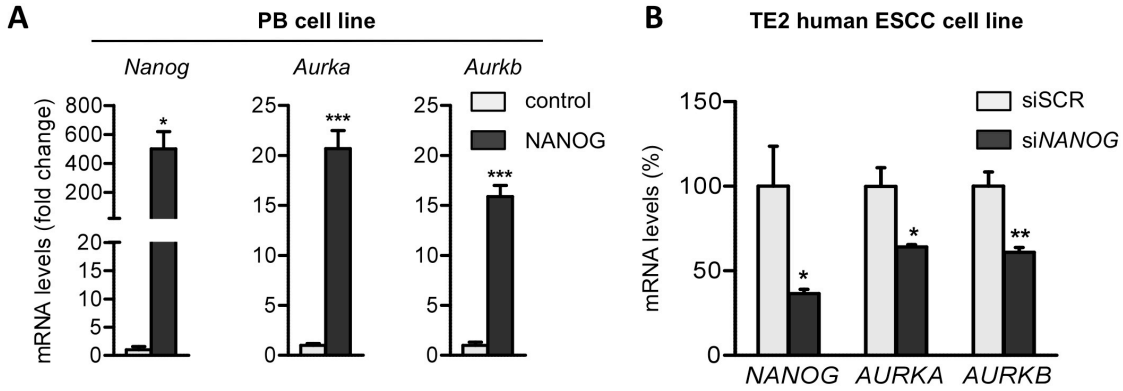


Figure 21. Nanog overexpression leads to upregulation of mitotic targets in papillomas-derived cells

(A) Relative mRNA levels of the indicated genes in PB cells transfected with a mouse NANOG expressing vector or mock transfected (control). Samples were analyzed 72 hr after transfection (n=3 technical replicates).

(B) Relative mRNA levels of the indicated genes in esophageal SCC (ESCC) TE2 cells transfected with pools of scrambled (si-scr) or anti-*NANOG* (si-*NANOG*) siRNAs. Samples were analyzed 48 h after transfection. Values correspond to mean ± s.e.m. (n=5 technical replicates).

Levels of mRNA were normalized to *Gapdh* and then compared to the control. Values correspond to mean ± s.e.m. Statistical significance was determined by the two-tailed Student's t test: **P*<0.05; ****P*<0.001.

Next, we wondered whether *Aurka* was a direct target of NANOG in epithelial cells. Inspection of previously published chromatin immunoprecipitation (ChIP) data of NANOG in mouse ESCs (Marson et al., 2008) uncovered a NANOG binding site ~7 kb upstream of the *Aurka* transcriptional start site.

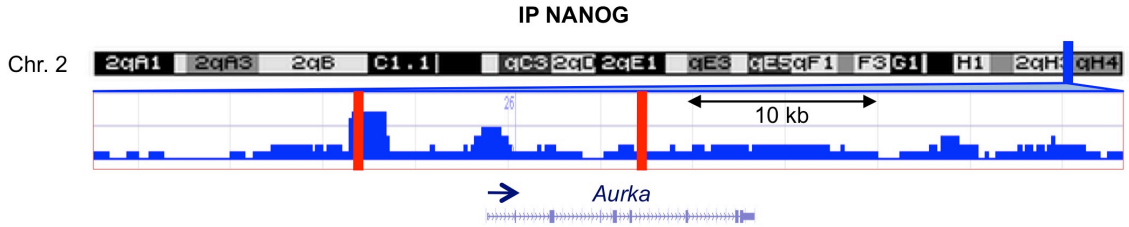


Figure 22. NANOG binding site in the *Aurka* promoter
Peaks of NANOG binding to the mouse *Aurka* gene determined by chromatin immunoprecipitation sequence (ChIP-seq). Top row, ideogram of mouse chromosome 2 with *Aurka* located in 2qH3 (blue bar). Middle row, blue histogram corresponds to the frequency of NANOG ChIP reads and red bars correspond to the regions chosen to be amplified in the ChIP experiments. The leftmost bar corresponds to the *Aurka* promoter and the rightmost bar to the intronic region. Bottom row, *Aurka* gene exon-intron structure

Based on this, we performed NANOG ChIP in WT and *Nanog*-KO ESCs and in primary mouse skin keratinocytes. For this, we took advantage of *p53*-null keratinocytes because of their robust proliferation and their higher basal levels of NANOG **(Figure 23A)**, the latter being in agreement with the negative regulation of *Nanog* expression by p53 (Lin et al., 2005). We confirmed that NANOG binds the *Aurka* promoter in ESCs. Importantly, NANOG also bound the *Aurka* promoter in keratinocytes **(Figure 23B)**, whereas it did not bind to an intronic *Aurka* region or to the *Oct4* promoter, which is not active in somatic tissues (Lengner et al., 2007). As a positive control, we confirmed that NANOG binds the *Oct4* promoter in ESCs **(Figure 23B)**.

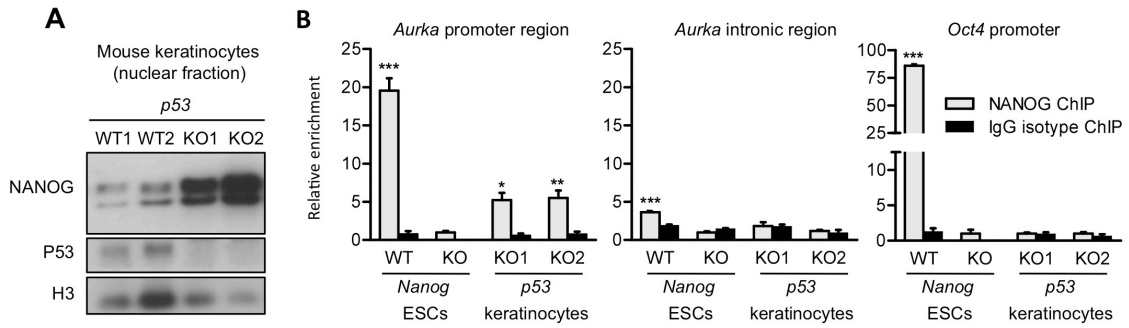


Figure 23. *Aurka* is a direct target of NANOG in epithelial cells
(A) Immunoblot of NANOG (using Millipore antibody) and p53 in nuclear extracts from primary mouse keratinocytes of the indicated genotypes. Histone H3 is used as loading control. We used two independent biological replicates, labeled WT1, WT2, KO1 and KO2, each from a different newborn mouse.
(B) Chromatin immunoprecipitation (ChIP) of NANOG in WT or *Nanog*-KO ESCs or in primary *p53*-KO keratinocytes. The relative recovery with respect to NANOG ChIP in *Nanog* KO ESCs is shown for *Aurka* promoter or intronic regions, or for the *Oct4* promoter. In the case of keratinocytes, we used two independent biological replicates, labelled KO1 and KO2, each from a different newborn mouse. Values correspond to mean ± s.e.m. of qPCR triplicates (*Aurka* promoter or intronic regions) or duplicates (*Oct4* promoter). Statistical significance vs. NANOG ChIP in *Nanog*-KO ESCs was determined by the two-tailed Student's t test: **P*<0.05; **P*<0.01; ****P*<0.001.

We conclude that NANOG is acting as a direct transcriptional activator of the mitotic kinase AURKA and this function is conserved in ESCs and in stratified epithelia.

1.5. NANOG binds to epithelial stem cell transcription factor ΔNp63α

To understand why NANOG is specifically causing an effect in stratified epithelia and not in other tissues, we hypothesized that it must bind to specific epithelial transcription factors that regulate its role. Throughout our results, we have observed that NANOG has the same expression pattern as P63. Moreover, recent results have identified SOX2 to bind to P63 where both share a transcriptional program in SCC cells (Watanabe et al., 2014a). We overexpressed NANOG and ΔNp63α, which is the predominant p63 isoform expressed in stratified epithelia and is essential for epithelial maintenance and epidermal morphogenesis (Mills et al., 1999; Yang et al., 1998), in U2OS cells. We found by co-immunoprecipitation (CoIP), in cells exogenously overexpressing both proteins, that NANOG can bind to ΔNp63α **(Figure 24A)**. To confirm this result using the endogenous proteins, we used ESCC cells, since ΔNp63α is known to be expressed in SCCs (Rocco et al., 2006; Yang et al., 1998). In TE12 cells (ESCC cell line) we confirmed endogenous NANOG binds to ΔNp63α **(Figure 24B)**. From these results we conclude NANOG can form a complex with ΔNp63α, which could mediate the specific role of NANOG in stratified epithelia.

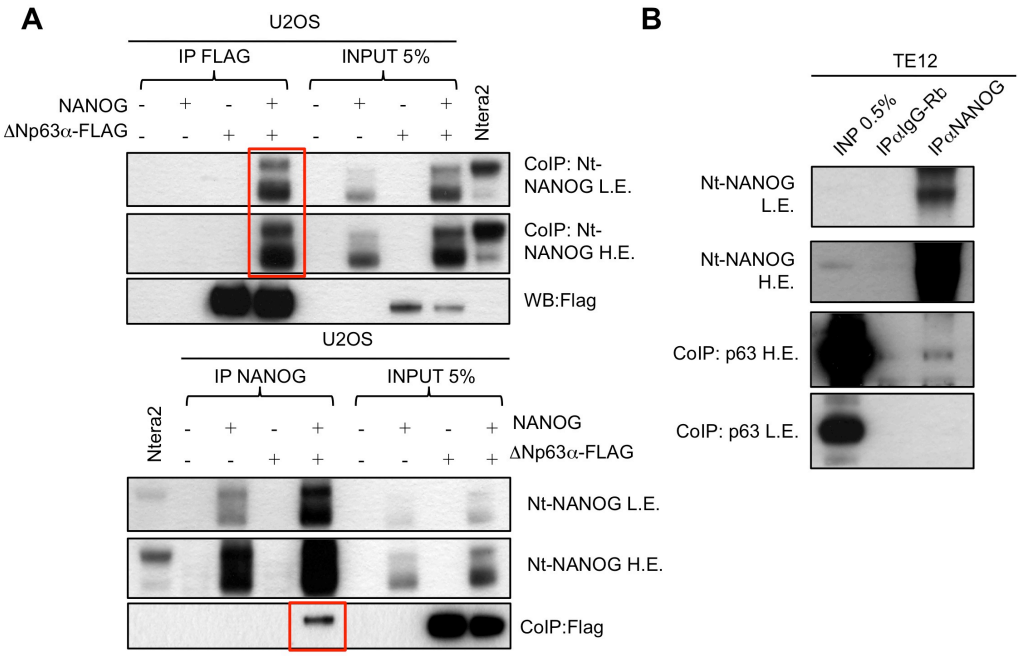


Figure 24. NANOG binds to epithelial stem cell marker ΔNp63a.

(A) Co-immunoprecipitation (CoIP) studies were carried out with lysates prepared from U2OS cells coexpressing the *NANOG* and ΔNp63α–FLAG constructs. For control purposes, U2OS cells were also transfected with *NANOG* or ΔNp63α–FLAG alone. Top: Proteins were immunoprecipitated with a C-terminal NANOG antibody and then immunoblotted using a FLAG monoclonal antibody. Bottom: In a reciprocal fashion, proteins were immunoprecipitated with the FLAG antibody and then immunoblotted using a N-terminal antibody. CoIP is indicated by a red box.

(B) CoIP was performed in ESCC TE12 cells that express endogenous *NANOG* and ΔNp63α. Proteins were immunoprecipitated with a N-terminal NANOG antibody and then immunoblotted using a p63 antibody. As a control, proteins were also immunoprecipitated using rabbit IgG isotype.

L.E. and H.E. refer to low and high exposition times respectively after ECL incubation.

All together our results demonstrate NANOG is functional in specific adult tissues, where its overexpression can activate a mitotic transcriptional program that leads to hyperplasia.

PART II
Role of NANOG in squamous cell carcinomas

2.1. NANOG is overexpressed in mouse and human squamous cell carcinomas

The mitogenic activity of NANOG in stratified epithelia (shown in Part I) prompted us to investigate the possible role of NANOG in squamous cell carcinomas. We began by exploring NANOG expression in a number of different mouse tumors. Interestingly, those tumors derived from stratified epithelia, such as forestomach, skin and vagina, were strongly positive for NANOG. Mammary and prostate tumors contained NANOG-positive cells, albeit of moderate intensity and restricted to a fraction of tumor cells. Tumors derived from NANOG-negative tissues, such as fibrosarcomas, intestinal adenomas, and hepatocellular carcinomas (HCC), were completely negative for NANOG **(Figure 25)**. These observations suggest that NANOG is strongly and widely expressed in cancers derived from stratified epithelia.

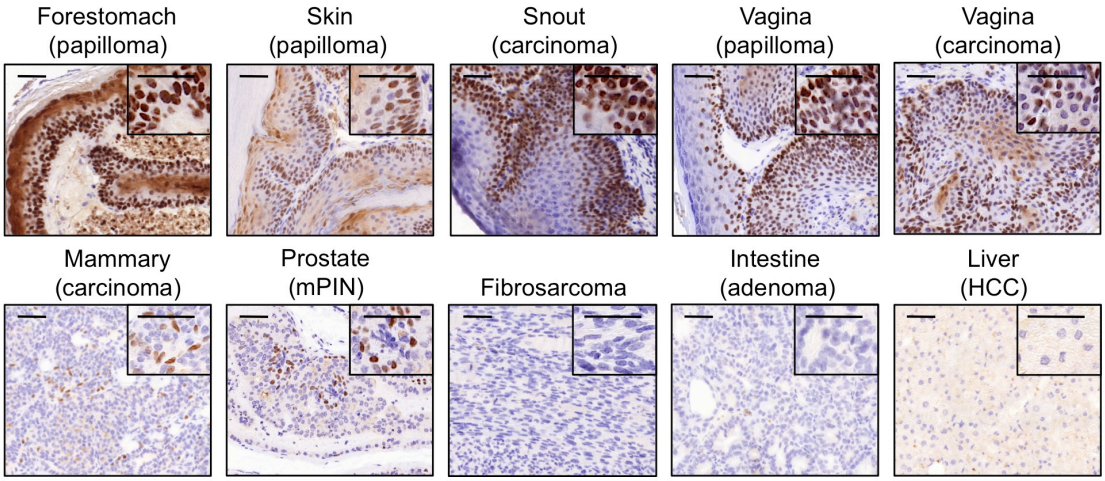


Figure 25. NANOG is overexpressed in tumors derived from stratified epithelia
NANOG IHC in archived mouse tumours, as indicated. mPIN: mouse prostate intraepithelial neoplasia; HCC: hepatocellular carcinoma. Bars correspond to 50 μm.

Based on the above, we decided to examine NANOG in human cancers derived from stratified epithelia. We focused on esophageal squamous cell carcinomas (ESCCs) and head and neck squamous cell carcinomas (HNSCCs) for which there is evidence of NANOG expression (Chiou et al., 2008; Du et al., 2012; Shimada et al., 2012). These studies were performed in collaboration with Manuel Rodriguez-Justo (University College Hospital, London, UK) who provided and analyzed ESCCs and with Marta Sanchez-Carbayo (CIC bioGUNE, Derio, Spain) who provided and analyzed HNSCCs. We confirmed that NANOG expression is common in ESCCs **(Figure 26A)** and HNSCCs **(Figure 26B)**. In fact, all the examined ESCCs (n=29) and HNSCCs (n=46) were positive for NANOG, often with very intense and widespread staining **(Figure 26A and 26B)**.

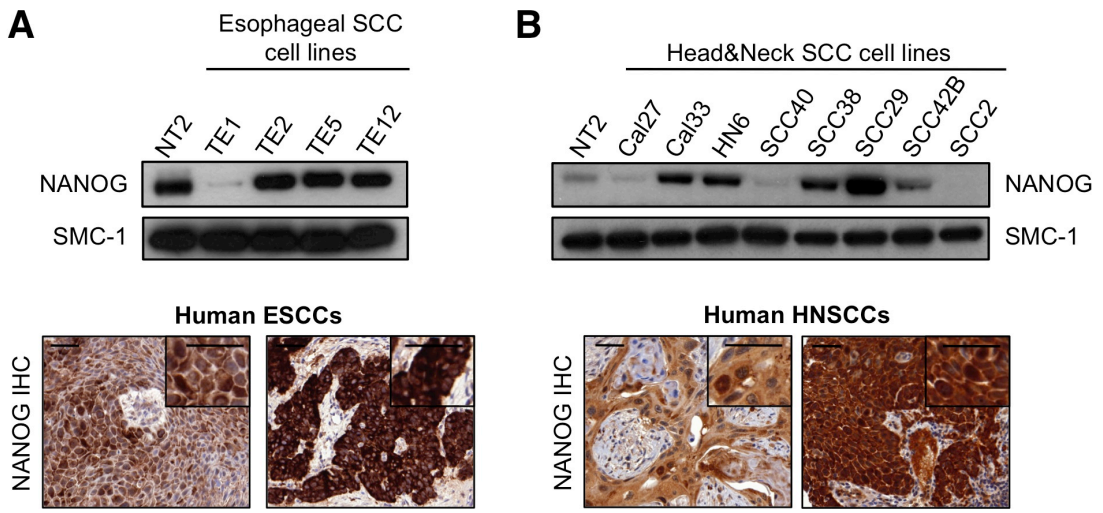


Figure 26. NANOG is expressed in human ESCC and HNSCC and in derived cell lines

(A) Top: Immunoblot of human NANOG in nuclear extracts from the indicated esophageal squamous cell carcinomas (ESCC) cell lines. Ntera2 (NT2) is used as a positive control for NANOG. SMC1 is used as a loading control. Bottom: IHC of NANOG and AURKA in two representative human ESCCs.

(B) Top: Immunoblot of human NANOG in nuclear extracts from the indicated head and neck squamous cell carcinomas (HNSCC) cell lines. Ntera2 (NT2) is used as a positive control for NANOG. SMC1 is used as a loading control. Bottom: IHC of NANOG and AURKA in two representative human HNSCCs.

Additionally, we also wanted to know if knockdown of NANOG in HNSCC cell lines has an effect in cell proliferation and cell cycle progression. As observed previously, many HNSCC cell lines express NANOG **(Figure 26B)**. We took two HNSCC cell lines and infected them with lentiviruses encoding shRNAs that target *NANOG* or scrambled sequences (which has no target in the genome) as a negative control. Immunoblot analysis showed knockdown of NANOG leads to decreased levels of AURKA protein **(Figure 27A)**. In agreement with this, *in vitro* 5-ethynyl-2′-deoxyuridine (EdU) proliferation assays indicate that silencing NANOG promotes a decrease of proliferation of HNSCC cells **(Figure 27B and 27C)**. Further analysis showed that NANOG knockdown also caused alterations in their cell cycle compared to the control **(Figure 27B and 27D)**.

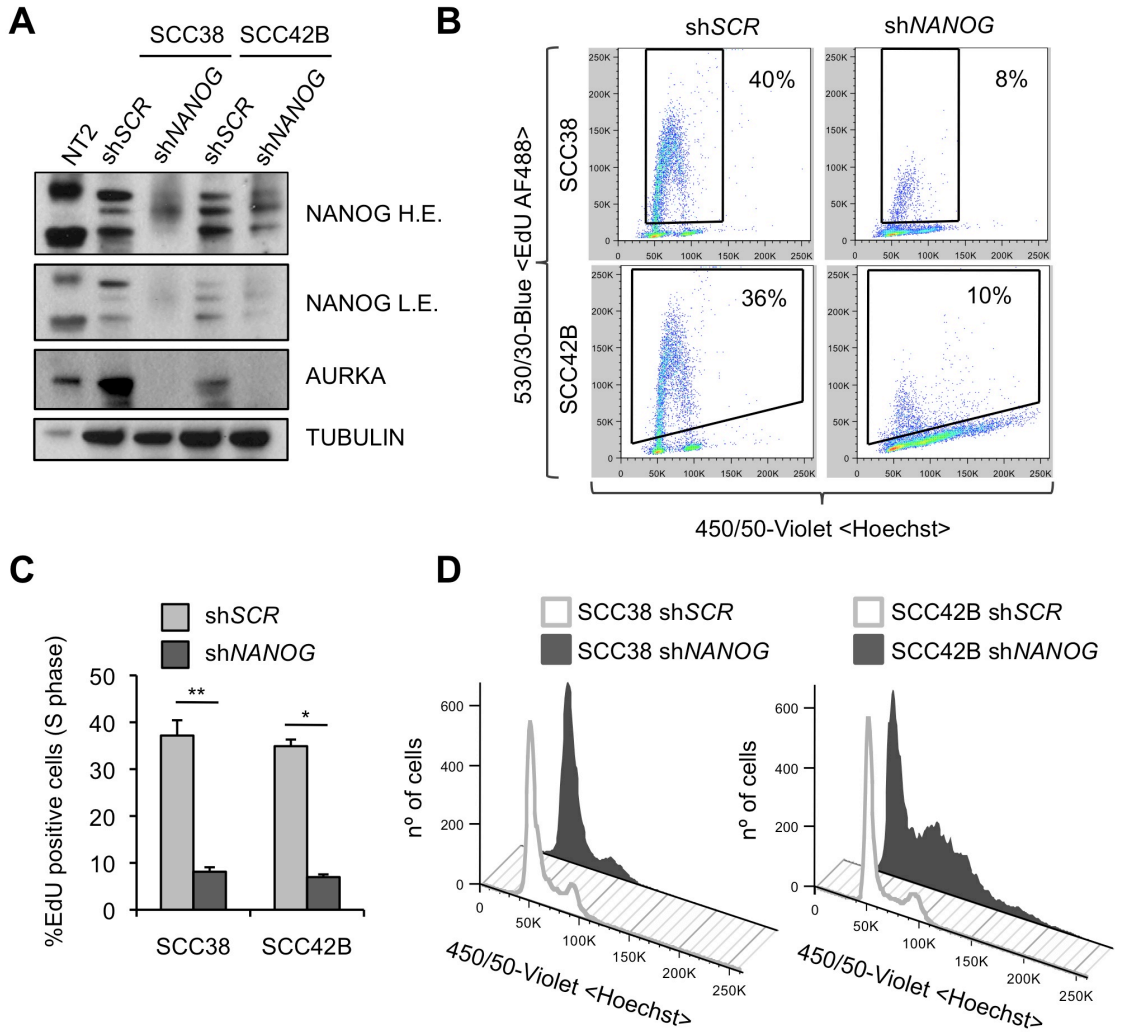


Figure 27. Knockdown of *NANOG* in HNSCC cells decreases proliferation and promotes cell cycle arrest

(A) Immunoblots of HNSCC cells (SCC38 and SCC42B) infected with sh*SCR* (control) or sh*NANOG* after selection with puromycin. NANOG knockdown promotes AURKA downregulation. TUBULIN is used as a loading control. L.E., low expostion after ECL incubation; H.E., high exposition after ECL incubation.

(B) HNSCC cells were treated as in A. After puromycin selection, cells were subject to 1 hr EdU-pulse, collected and processed using the Click-It EdU kit (Invitrogen), together with Hoechst staining, and analyzed by flow cytometry. EdU positive cells correspond to the proliferating cells (S phase). Both cell lines upon *NANOG* knockdown showed cell cycle arrest and decreased proliferation.

(C) Quantification of EdU positive cells after flow cytometry analysis in SCC38 and SCC42B cells treated as in B. Values correspond to mean ± SD in SCC38 (n=3 biological replicates) and SCC42B (n=2 biological replicates). Statistical significance was determined by the two-tailed Student's t test: (*) $P<0.05$; (**) $P<0.01$; non significant (n.s.).

(D) Cell cycle analysis using Hoechst staining of DNA content of HNSCC cells treated as in B.

We conclude that NANOG expression is prevalent in many types of squamous cell carcinoma and this correlates also with its function in stratified epithelia. These observations reinforce the concept that NANOG could have an effect not only in stratified epithelia, but also in their derived cancers.

2.2. NANOG promotes proliferation in the epidermis

Based on our previous observations, our next objective was to address the *in vivo* role of NANOG in oncogenesis. *Nanog* ubiquitous overexpression showed a decreased survival of mice, where NANOG-overexpressing mice died after a few weeks and is not a suitable model

for long-term carcinogenesis assays. For this reason, we decided to use a tissue specific doxycycline-inducible *Nanog* transgenic mouse model. In this case, instead of using the *Rosa26::rtTA* mouse, we chose another available mouse model (*K5-rtTA*) kindly provided by J. Silvio Gutkind (NIH, Bethesda, USA), which has the bovine keratin 5 (K5) promoter directing the expression of the reverse tetracycline transactivator gene (rtTA) (Vitale-Cross et al., 2004) **(Figure 28A)**. The combination of these transgenes (*K5-rtTA;Col1a1::tetO-Nanog*) provides a Tet-ON tool that allows the inducible expression of *Nanog* in the basal layer of the skin, which includes the stem cell compartments (Fuchs, 2009). We first examined whether transgenic *Nanog* was specifically expressed in the basal layer of stratified epithelium. We treated adult *K5-rtTA*$^{tg/.}$;*Col1a1*$^{tetO\text{-}Nanog/+}$ (from now on abbreviated as TG) and their control littermates lacking the transactivator *Col1a1*$^{tetO\text{-}Nanog/+}$ (from now on abbreviated as CTR) with doxycycline in their drinking water (2 mg/ml) for a period of 48 hours. We observed *Nanog* expression by qRT-PCR selectively in stratified epithelia-containing tissues, such as, forestomach, tongue, tail and back skin and esophagus. As a negative control, we could not detect expression of transgenic *Nanog* in the liver and the small intestine (non-stratified epithelial tissues) **(Figure 28B)**. We also performed immunohistochemistry (IHC) to confirm NANOG overexpression in the basal layer of the skin of TG mice **(Figure 28C)**. After 48-hour transgenic *Nanog* overexpression we could not detect histological alterations in the skin **(Figure 28C)**.

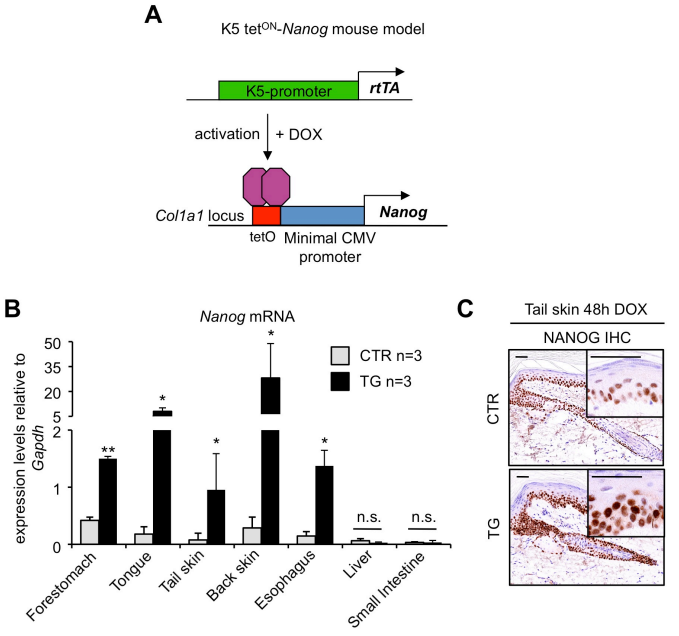


Figure 28. K5 *Nanog*-inducible mouse model

(A) Schematic representation of the *K5-rtTA;Col1a1::tetO-Nanog* mouse model. The reverse transactivator (rtTA) is under the control of the bovine cytokeratin 5 (K5) promoter. A cassette containing the *Nanog*-cDNA is under the control of the doxycycline (DOX)-responsive promoter (*tetO*), which was inserted downstream of the *Col1a1* locus. tetO, tetracycline/ doxycycline-responsive operator. Arrows indicate transcriptional start sites.

(B) Relative *Nanog* mRNA levels in the indicated organs determined by qRT-PCR. *Nanog*-inducible mice (n=3) containing both transgenes (*Col1a1*$^{tetO\text{-}Nanog/+}$;*K5-rtTA*$^{tg/.}$; abbreviated as TG) and control mice (n=3) lacking the transactivator (*Col1a1*$^{tetO\text{-}Nanog/+}$; abbreviated as CTR) were analyzed 48 hr after administration of doxycycline (DOX) (2mg/ml) in the drinking water.

(C) Immunohistochemistry (IHC) for NANOG of paraffin-embedded sections of tail skin from CTR and TG mice treated as indicated in (B). Two magnifications are shown for each tissue (bars correspond to 50 µm).

To study the long-term effects of *Nanog* expression, mice were subjected to continuous doxycycline administration via the drinking water during 9 months. Transgenic expression of NANOG was clearly detectable in the basal layer of the skin after 9 months, but no alterations were observed in its tissue architecture (**Figure 29**). Esophagus and forestomach showed mild

hyperplasia in TG mice in agreement with our previous results, but this does not lead to decreased survival (**Figure 29**).

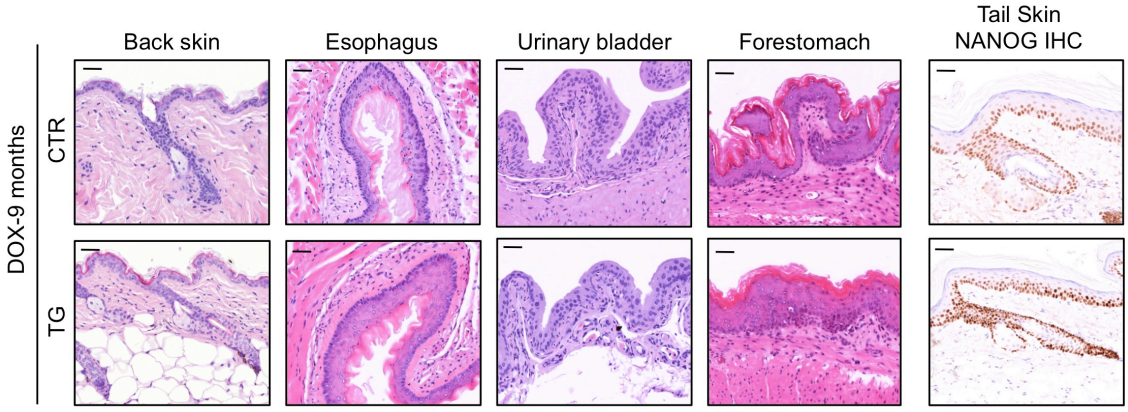


Figure 29. Long-term *Nanog* overexpression in K5 *Nanog*-inducible mice
Representative hematoxylin and eosin (H&E) staining of CTR and TG mice after 9 month-doxycycline treatment of epithelial tissues: back skin, esophagus, urinary bladder and forestomach. Immunohistochemistry (IHC) for NANOG of paraffin-embedded sections of tail skin from CTR and TG mice treated with doxycycline during 9 months (right). Bars correspond to 50 µm.

To investigate the effect of NANOG in the context of mitogenic and inflammatory stimulation, we topically treated the tail skin of CTR and TG mice with 12-*O*-tetradecanoylphorbol 13-acetate (TPA). In mice overexpressing transgenic *Nanog* **(Figure 30A)**, four successive applications of TPA at 48-hour intervals resulted in augmented epidermal thickness compared to CTR mice **(Figure 30B)**. This effect was accompanied by a clear increase in Ki67-positive proliferating keratinocytes in the basal and suprabasal layers of TG mice compared to CTR ones, as well as a decrease of the differentiation marker LORICRIN **(Figure 30C)**. Therefore, NANOG has the capacity to induce epidermal hyperproliferation under conditions of mitogenic stimulation.

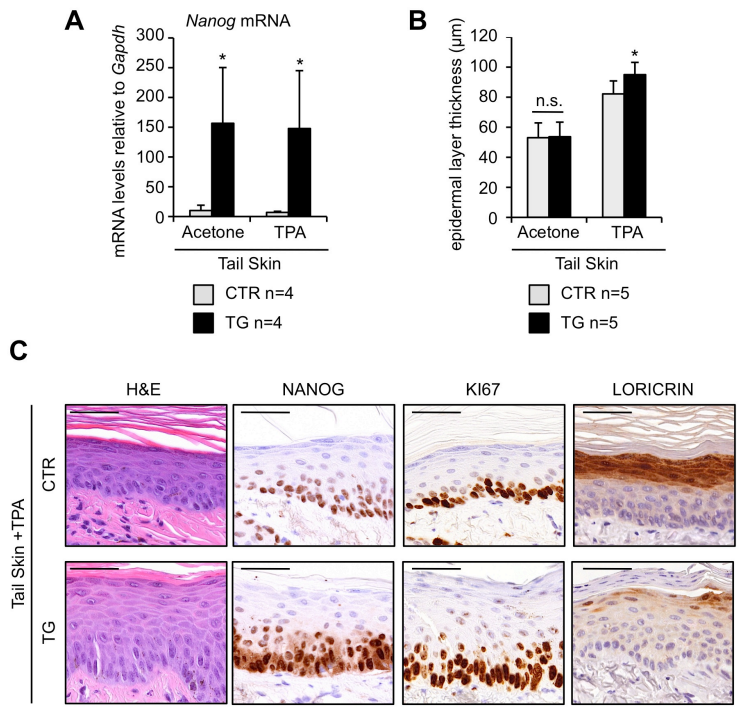


Figure 30. *Nanog* overexpression promotes epidermal hyperproliferation upon TPA application
(A) Relative *Nanog* mRNA levels in CTR and TG tail skin topically treated 4 times, at 48 hr intervals, with 12-*O*-tetradecanoylphorbol 13-acetate (TPA) or acetone as a control. Mice (n=4) were exposed to doxycycline (2mg/ml) in their drinking water 48 hr prior to TPA treatment and throughout the remaining protocol.
(B) Epidermal thickness (basal and suprabasal layers) of CTR and TG mice (n=5) treated as in (A).
(C) Representative hematoxylin and eosin (H&E) staining and immunohistochemistry (IHC) for NANOG, KI67 and LORICRIN of TPA-treated CTR and TG mice as indicated in (D) (bars correspond to 50 µm).
In (A) and (B) values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: (*) $P<0.05$; (**) $P<0.01$.

2.3. NANOG promotes squamous cell carcinoma (SCC) formation

As shown previously, NANOG is overexpressed in squamous cell carcinomas. To address whether high levels of NANOG can contribute to tumorigenesis, we induced skin tumors using the classical DMBA/TPA two-stage chemical carcinogenesis protocol (Abel et al., 2009). We used CTR and TG mice in a C57BL/6 background, which is known to be resistant to carcinoma formation (Hennings et al., 1993), and administered doxycycline throughout the entire protocol **(Figure 31A)**. There were no significant differences in the kinetics of papilloma formation **(Figure 31B)**, or in the total number **(Figure 31C)** or size **(Figure 31D)** of papillomas at the end of the experiment.

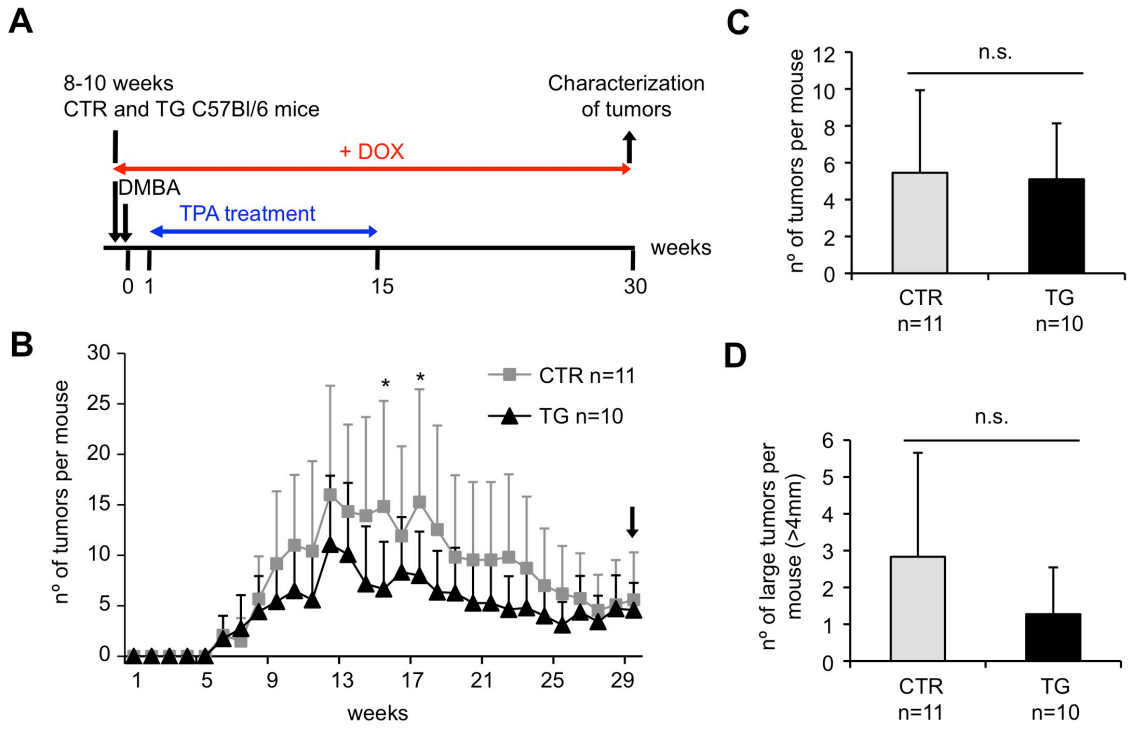


Figure 31. Chemical carcinogenesis with DMBA/TPA

(A) Experimental layout for two-stage chemical-induced carcinogenesis. CTR and TG mice (see Figure 1A) were treated with doxycycline (DOX) in their drinking water (2 mg/ml) during the entire protocol. DMBA was applied 48 hr after initiation of DOX treatment. TPA application started the following week, twice a week during 15 weeks. Tumors were counted and measured weekly. Histological analysis was performed at the humane endpoint or at the end of the experiment (30 weeks).

(B) Average number of tumors per mouse during the treatment. Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined for each time point by two-tailed Student't test: (*) $p<0.05$.

(C and D) Average number of total tumors (C) and large tumors (>4mm) (D) per mouse. Quantification was performed at the humane endpoint or at the end of the experiment (30 weeks). Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: n.s. non significant

Interestingly, at around 20 weeks of treatment, tumors resembling squamous cell carcinomas (SCCs) started to appear exclusively in the TG mice **(Figure 32A)**. In many cases (~40%), TG mice had to be sacrificed before the finalization of the observation period because carcinomas were larger than 1.5 cm (humane endpoint) **(Figure 32B)**. All tumors were characterized at the end of the protocol or at the humane endpoint. As expected for the C57BL/6 background, all CTR tumors were papillomas (benign). In contrast, a significant fraction of tumors (>15%) from

the TG mice had progressed to SCCs **(Figures 32C and 32D)**. Moreover, some of these SCCs presented areas with spindle component and in one case metastasis to a lymph node **(Figure 32E)**.

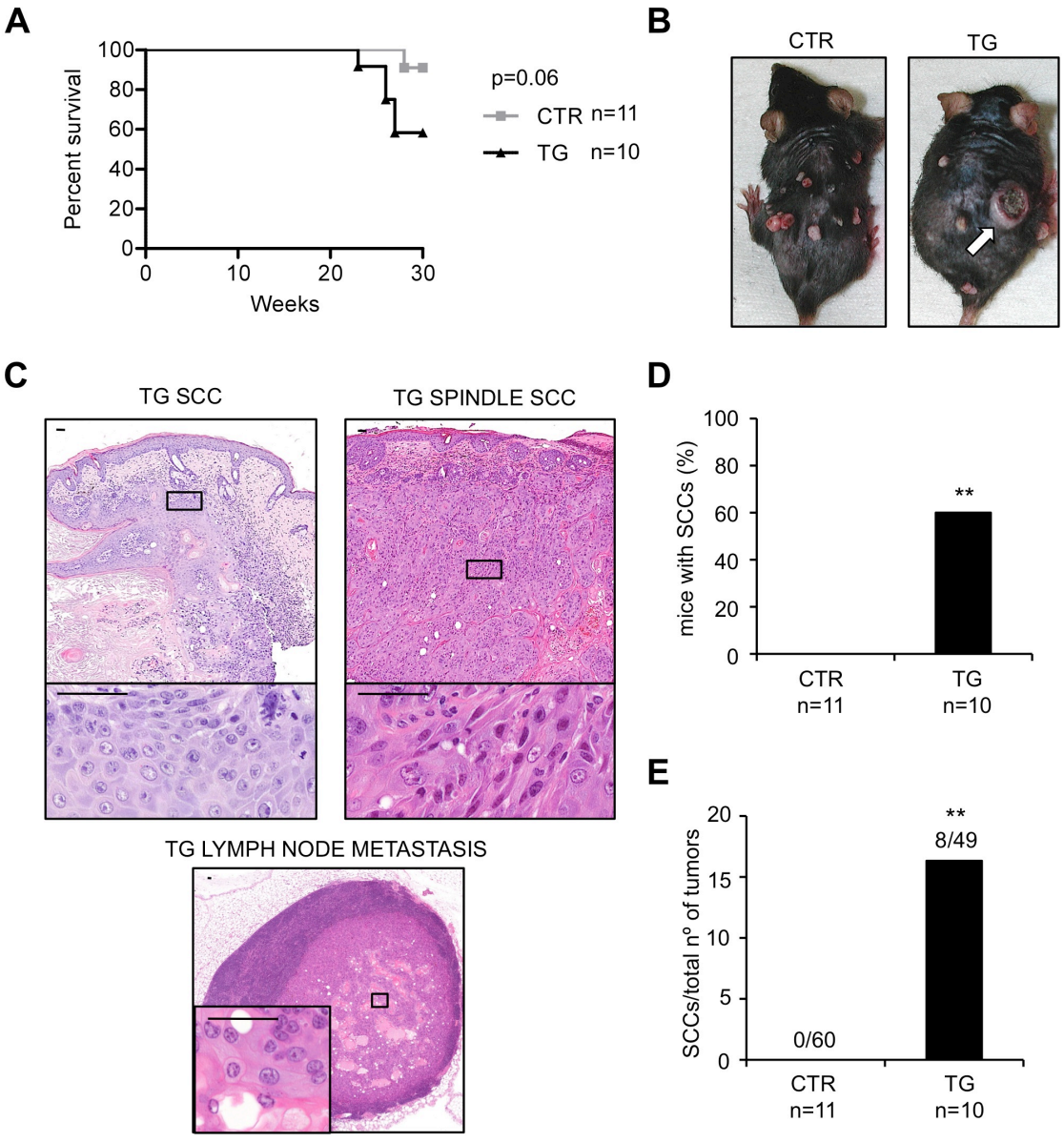


Figure 32. ***Nanog*** **overexpression promotes epidermal squamous cell carcinoma**

(A) Survival curve of CTR and TG mice upon DMBA/TPA treatment and continuous treatment with doxycycline. Statistical significance was determined by the Long-rank (Mantel-Cox) test. (n), number of mice.

(B) Representative images of DMBA/TPA treated mice. Arrow indicates a squamous cell carcinoma.

(C) Hematoxylin and eosin (H&E) staining of TG carcinomas and lymph node metastasis. Two magnifications are shown for each tissue (bars correspond to 50 µm).

(D) Percentage of CTR and TG mice with squamous cell carcinomas (SCCs) at the end of the treatment. Statistical significance was determined by the two-tailed Student's t test: (**) $p<0.01$.

(E) Conversion rate showing the number of SCCs relative to the total number of tumors. Statistical significance was assessed by the Fisher's exact test: (**) $p<0.01$.

We confirmed NANOG overexpression in the TG mice skin, papillomas and carcinomas **(Figure 33A and 33B)**. We also noticed that the endogenous *Nanog* gene was upregulated in TG mice **(Figure 33B)**, suggesting that transgenic *Nanog* may upregulate the endogenous

Nanog gene. Interestingly, carcinomas showed the highest levels of *Nanog* (transgenic and endogenous), suggesting an association with malignancy. To explore this further, we analyzed NANOG protein levels in different cell lines previously obtained from these types of tumors kindly provided by Miguel Quintanilla (Instituto de Investigaciones Biomédicas "Alberto Sols", Madrid). Spindle SCC-derived cell lines (CarB, CarC and MSC11A5) expressed the highest levels of NANOG, compared to cell lines derived from SCCs, papillomas or immortalized keratinocytes **(Figure 33C)**.

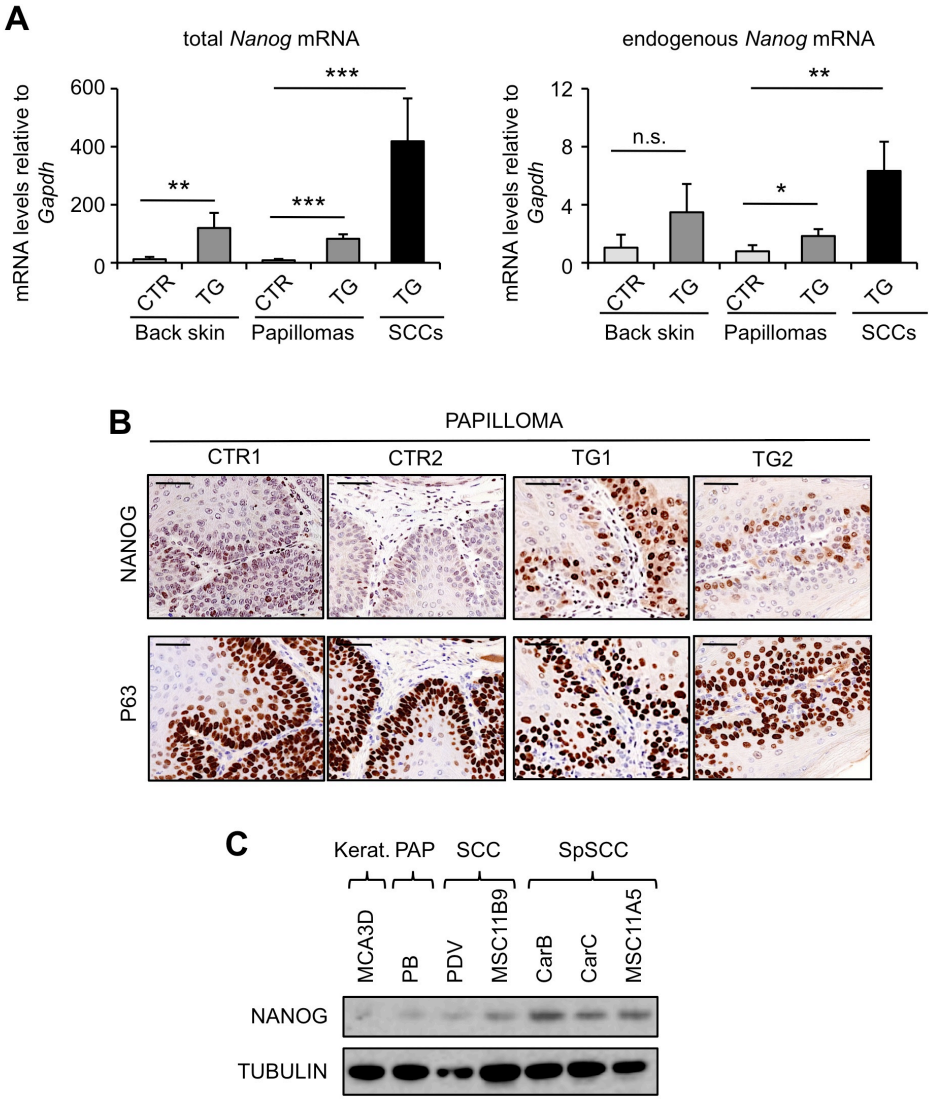


Figure 33. Nanog expression levels in back skin, papillomas and SCCs

(A) Relative total *Nanog* mRNA levels (left) and endogenous *Nanog* mRNA levels (right) in CTR (n=4) and TG (n=4) back skin; CTR (n=4) and TG (n=3) papillomas; and TG (n=3) SCCs. mRNA levels were normalized by *Gapdh* levels. Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: n.s. non significant; (*)$P<0.05$; (**) $P<0.01$; (***) $P<0.001$

(B) IHC of NANOG and P63 in representative CTR and TG papillomas. Bars correspond to 50 µm.

(C) Immunoblots of NANOG using total lysates from the indicated cell lines. TUBULIN was used as a loading control. Kerat., immortalized keratinocytes; PAP, papilloma; SCC, squamous cell carcinoma; SpSCC, spindle SCC.

Together, these results demonstrate that NANOG can promote malignant progression to squamous cell carcinomas.

2.4. NANOG induces EMT targets in skin papillomas

To obtain mechanistic insight on how NANOG promotes skin tumorigenesis, we performed with the help of Orlando Dominguez (CNIO), Osvaldo Graña (CNIO) and Gonzalo Gómez López (CNIO), RNA-seq analysis of CTR (n=4) and TG (n=3) papillomas to capture early events during *Nanog*-driven tumorigenesis. This analysis revealed a small number of differentially-expressed genes (98 genes; see **Table S1).** Interestingly, genes associated to epithelial-mesenchymal transition (EMT) were remarkably over-represented in *Nanog*-papillomas, including *Prrx1*, a master driver of EMT (Ocana et al., 2012). This was confirmed by gene set enrichment analysis (GSEA) of a number of EMT signatures (Blick et al., 2010; Jechlinger et al., 2003; Taube et al., 2010) **(Figure 34)**.

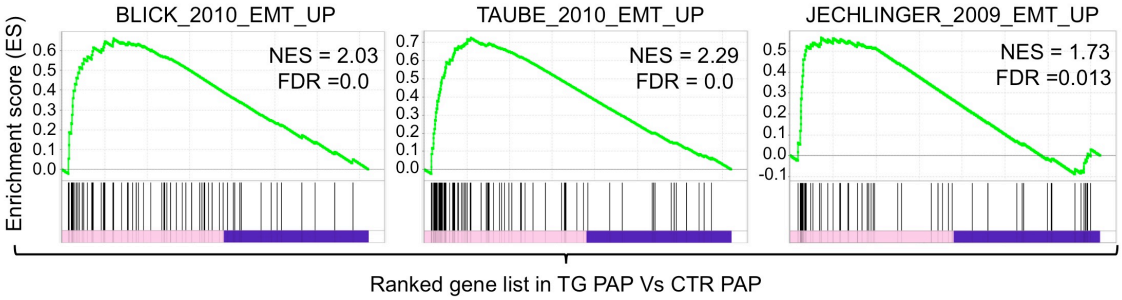


Figure 34. *Nanog* overexpression in papillomas induces an EMT signature
Gene set enrichment data (GSEA) showing the enrichment of three published EMT gene signatures (Blick et al., 2010; Jechlinger et al., 2003; Taube et al., 2010) in TG papillomas as compared with CTR papillomas. NES: normalized enrichment score. The false discovery rate (FDR; q-value) is indicated.

We validated the upregulation of EMT drivers and markers by qRT-PCR **(Figure 35A)** and by immunohistochemistry **(Figure 35B)** in *Nanog*-driven papillomas and carcinomas. Of note, downregulation of E-cadherin was also observed **(Figure 35A and 35B)**.

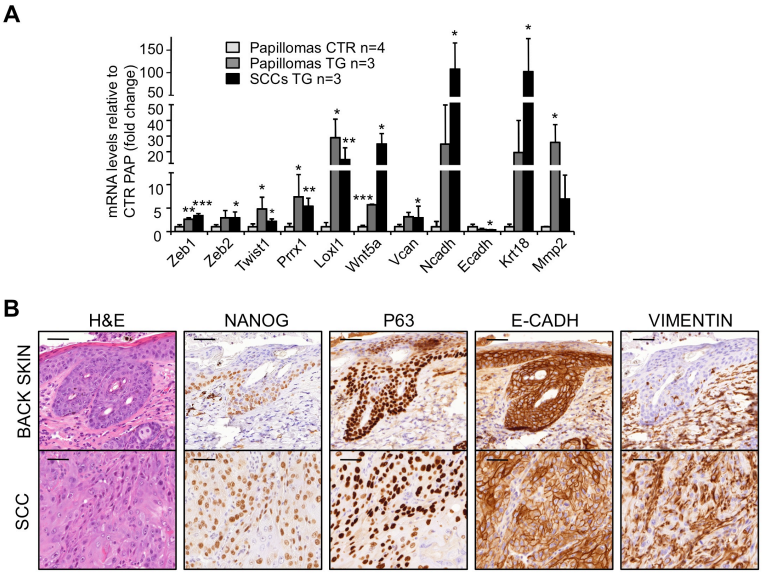


Figure 35. *Nanog* overexpression upregulates EMT targets in vivo.
(A) Relative gene expression in CTR and TG papillomas and TG SCCs of stemness related genes analyzed by qRT-PCR. mRNA levels were normalized by *Gapdh* levels and, then compared to levels of CTR papillomas (shown as fold change). Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test relative to CTR papillomas: (*) $P<0.05$; (**) $P<0.01$; (***) $P<0.001$.
(B) Representative hematoxylin and eosin (H&E) staining of a TG SCC. Top corresponds to magnified back skin and bottom to magnified invasive carcinoma. Serial sections were stained by IHC for NANOG, P63, E-CADHERIN and VIMENTIN (bars correspond to 50 µm).

Based on this, we wanted to know whether NANOG had the ability to initiate an EMT program in primary keratinocytes. For this, we extracted keratinocytes from CTR and TG neonates and added doxycycline to their culture medium for 48 hours. Interestingly, NANOG was able to upregulate a subset of EMT mediators in primary keratinocytes as soon as 48 hours after induction (**Figure 36A and 36B)**.

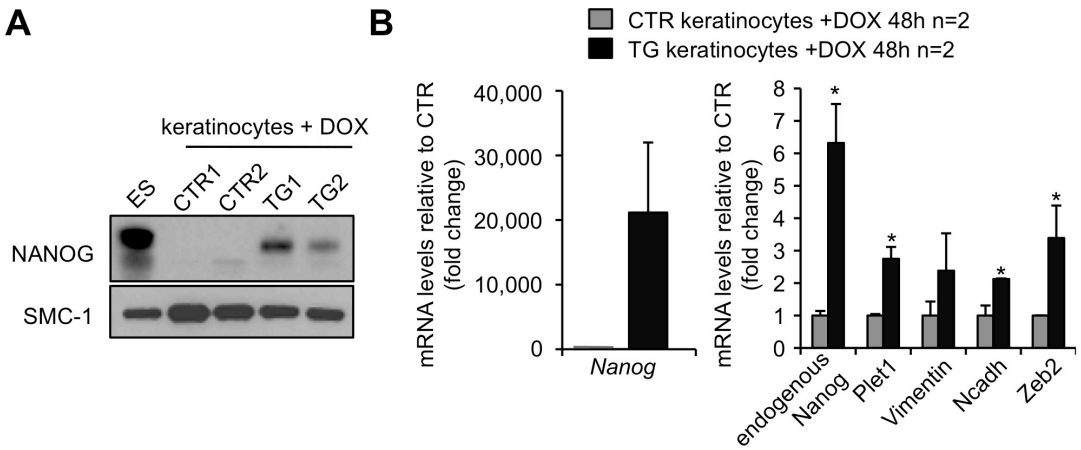


Figure 36. *Nanog* overexpression in primary keratinocytes upregulates EMT markers

(A) Immunoblots of NANOG using nuclear lysates from CTR and TG keratinocytes after 48 hr doxycycline treatment. SMC-1 was used as a loading control. Mouse embryonic stem cells (ES) were used as a positive control.

(B) Relative mRNA levels of the indicated genes in primary keratinocytes extracted from CTR and TG newborn pups (n=2) and treated during 48 hr with doxycycline. mRNA levels were normalized by *Gapdh* levels and then compared to levels of CTR keratinocytes (shown as fold change).

MicroRNAs have been also implicated as important regulators of EMT and malignancy, being particularly important miR-21, as a positive inducer, and miR-200 and miR-34 families, as negative regulators (Bornachea et al., 2012; Cufi et al., 2013; Dykxhoorn, 2010; Garzon et al., 2010; Ma et al., 2011; Medina et al., 2010). In collaboration with the group of Jesus M. Paramio (CIEMAT, Madrid), we examined the expression of these miRNAs in normal skin, papillomas and carcinomas. Interestingly, we found that oncomiR miR-21 is significantly upregulated by NANOG in back skin, papillomas and carcinomas **(Figure 37)**.

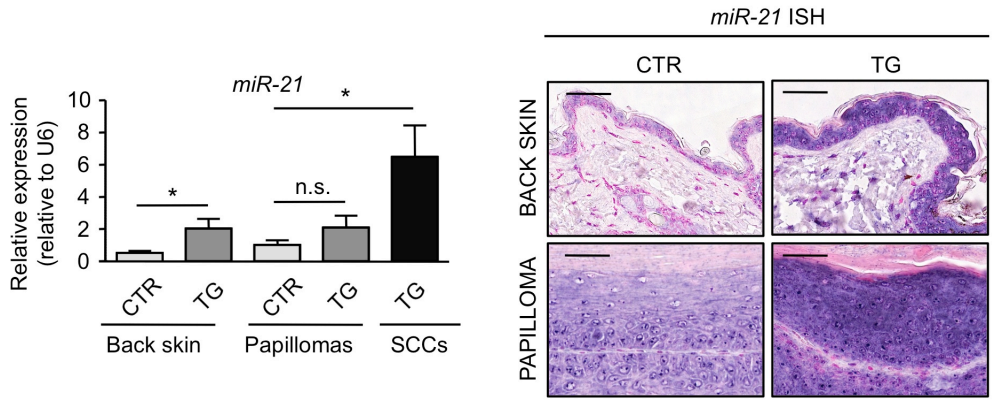


Figure 37. *Nanog* overexpression *in vivo* upregulates expression of *miR-21*

Left: Relative *miR-21* levels in CTR (n=4) and TG (n=4) back skin; CTR (n=4) and TG (n=3) papillomas; and TG (n=3) SCCs. Expression levels were normalized to U6. Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test: n.s., non significant; (*) $P<0.05$. Right: *In situ* hybrdization (ISH) of *miR-21* in representative CTR and TG back skin and papillomas (bars correspond to 50 µm).

In contrast, miRNAs from the miR-200 and miR-34 families were essentially unchanged between control and *Nanog* transgenic samples (**Figure 38)**.

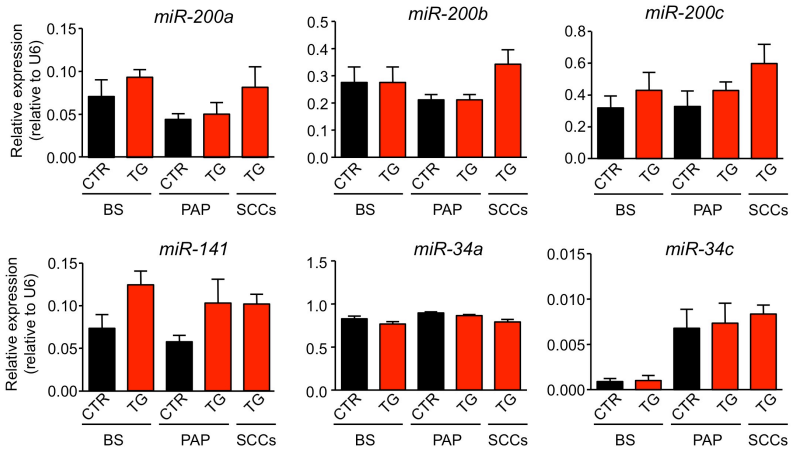


Figure 38. Nanog overexpression *in vivo* does not alter expression of miR-34 and miR-200 families. Relative expression levels of indicated microRNAs in CTR (n=4) and TG (n=4) back skin; CTR (n=4) and TG (n=3) papillomas; and TG (n=3) SCCs. Expression levels were normalized to U6. Values correspond to mean ± SD of the indicated number of mice (n). No statistically significant differences were found using the two-tailed Student's t test.

Importantly, NANOG in association to pSTAT3 has been shown to bind the promoter of *miR-21* and upregulate its expression in head and neck SCC cell lines (Bourguignon et al., 2012a). In this regard, we found that NANOG-overexpressing papillomas also express high levels of pSTAT3 **(Figure 39A)**, thereby lending support to the above-mentioned link between NANOG and pSTAT3 in SCCs. MicroRNA miR-21 has been reported to function through a variety of mechanisms, including upregulation of pAKT and pERK (Bornachea et al., 2012; Hatley et al., 2010; Ma et al., 2011). In agreement with this, we observed increased pAKT **(Figure 39A)** and pERK **(Figure 37B)** in NANOG-driven tumors.

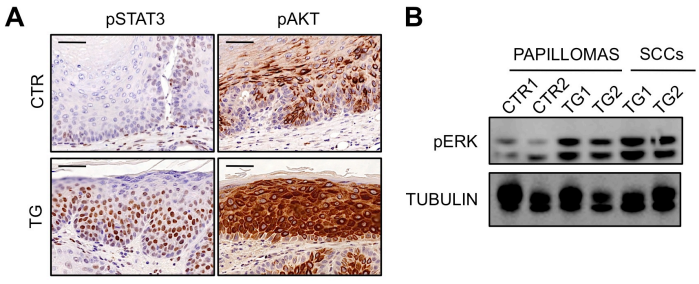


Figure 39. *Nanog* overexpressing papillomas and carcinomas have increased levels of pSTAT3, pAKT and pERK.
(A) IHC of phospho-STAT3 (Tyr705) and phospho-AKT (Ser473) in representative CTR and TG papillomas (bars correspond to 50 µm).
(B) Immunoblots of phospho-ERK (Thr202/Tyr204) in lysates from back skin, papillomas and SCCs from CTR and TG mice. TUBULIN was used as a loading control.

Together, these observations indicate that NANOG drives SCC tumorigenesis in association with the upregulation of EMT drivers, notably including *Prrx1* and *miR-21*.

2.5. NANOG induces stemness in skin papillomas

Examination of the list of differentially expressed genes in CTR and TG papillomas also revealed a high frequency of genes previously related to stemness **(Table S1)**. Gene set enrichment analysis (GSEA) on a cancer stem cell signature from breast cancer (Gupta et al., 2009) confirmed this idea **(Figure 40A)**. Several stemness-associated genes were validated by

qRT-PCR, including *Cxcl12/Sdf-1, Pecam1/CD31, Plet1, Lgr5, Lrg1, Pdgfra* and *H19*, which were all significantly upregulated in TG papillomas and carcinomas **(Figure 40B)**.

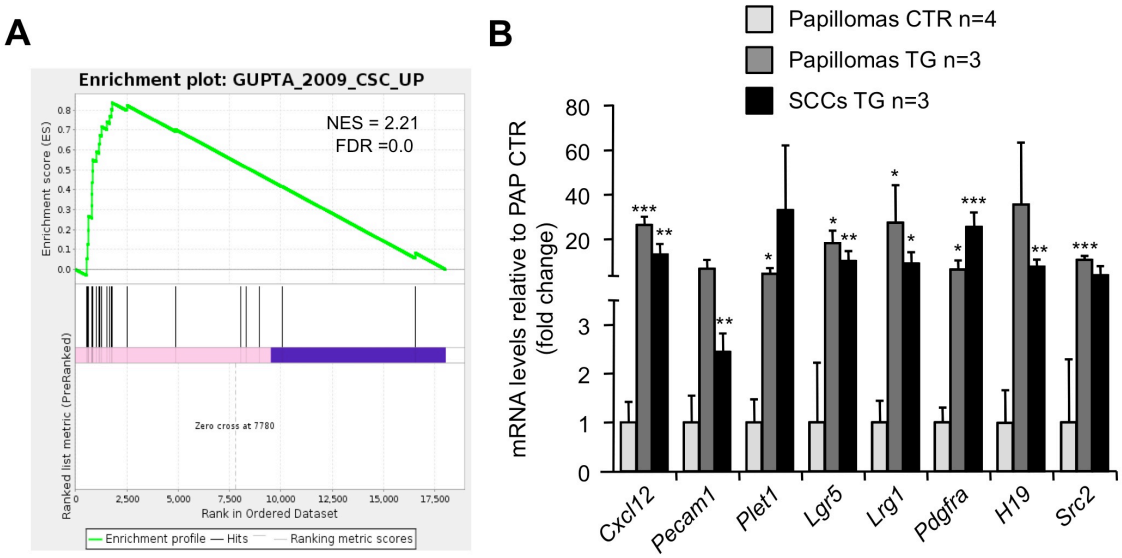


Figure 40. NANOG induces stemness in skin papillomas

(A) Gene set enrichment data (GSEA) showing the enrichment of a published breast cancer stem cell signature (Gupta et al., 2009) in TG papillomas as compared with CTR papillomas. NES: normalized enrichment score. The false discovery rate (FDR, q-value) is indicated.

(B) Relative gene expression in CTR and TG papillomas and TG SCCs of stemness related genes analyzed by qRT-PCR. mRNA levels were normalized by *Gapdh* levels and, then compared to levels of CTR papillomas (shown as fold change). Values correspond to mean ± SD of the indicated number of mice (n). Statistical significance was determined by the two-tailed Student's t test relative to CTR papillomas: (*) $P<0.05$; (**) $P<0.01$; (***) $P<0.001$.

2.6. NANOG directly activates EMT and stemness genes

NANOG is a well-known transcription factor in embryonic stem cells where it binds to a large amount of promoter regions (Loh et al., 2006). We took advantage of the ENCODE project (Rosenbloom et al., 2013) to examine ChIP-seq data of NANOG in the above-described EMT and stemness genes upregulated in NANOG-driven papillomas. All the examined genes had NANOG binding sites in human embryonic stem cells according to the ENCODE data (Rosenbloom et al., 2013) and these sites were confirmed by us using human embryonic carcinoma NTERA2 cells (**Figure 41A**). Interestingly, many of these NANOG binding sites were also occupied by endogeneous NANOG in human esophageal SCC cell line TE2 **(Figure 41B)**. In particular, we observed direct binding of endogenous NANOG to EMT mediators *MIR21*, *TWIST1, PRRX1* and *ZEB2*, as well as, to stemness associated genes *LGR5, PDGFRA, CXCL12* and *PITX2* **(Figure 41B)**.

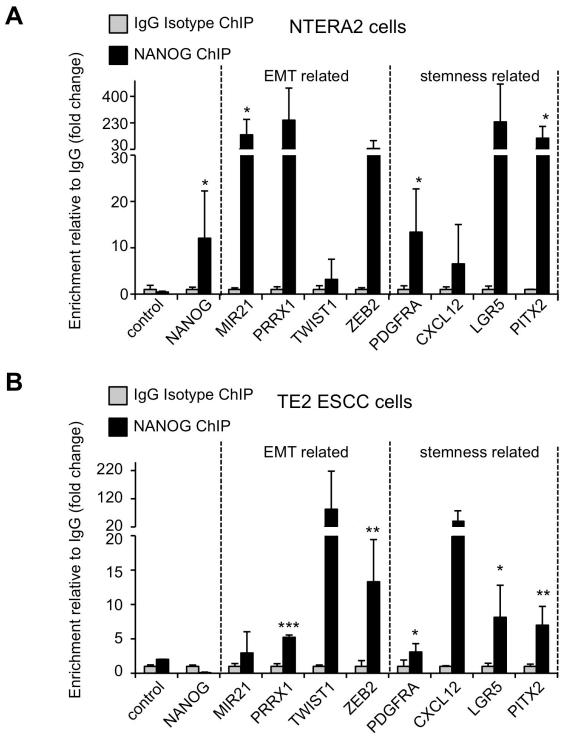


Figure 41. NANOG directly regulates EMT and stemness genes

(A) Chromatin immunoprecipitation (ChIP) of NANOG and goat IgG in NTERA2 teratocarcinoma cells.

(B) Chromatin immunoprecipitation (ChIP) of NANOG and goat IgG in TE2 esophageal SCC (ESCC) cells.

Values were first normalized to the input values, and then compared to levels of the IgG control (shown as fold change). Values correspond to mean ± SD of two independent biological replicates (n=2). Control, genomic region that does not bind NANOG (Gifford et al., 2013). Statistical significance was determined by the two-tailed Student's t test: (*) $P<0.05$. (*) $P<0.05$; (**).

2.7. NANOG promotes EMT in a cell-autonomous manner

To investigate if NANOG overexpression in keratinocytes could recapitulate the observed *in vivo* phenotype, we infected human immortalized HaCat keratinocytes with a lentivirus encoding NANOG under the CMV promoter, as well as a control empty lentivirus. After puromycin selection, we confirmed NANOG overexpression by immunoblotting **(Figure 42A)**, as well as by qRT-PCR **(Figure 42B)**. By immunofluorescence we confirmed that exogenous NANOG correctly localized to the nucleus **(Figure 42C)**. Additionally, we also confirmed upregulation of some EMT-inducing factors, such as *TWIST1* and *ZEB1* (**Figure 42B**). To examine the functional consequences of NANOG overexpression, we focused on some properties associated to EMT. In particular, NANOG overexpressing HaCat cells had a more pronounced migration capacity **(Figure 42D)**. Also, EMT is known to decrease the proliferative capacity of cells (Nieto and Cano, 2012) and, consistent with this, we observed that NANOG overexpressing HaCat cells had a decreased proliferation rate **(Figure 42E)**.

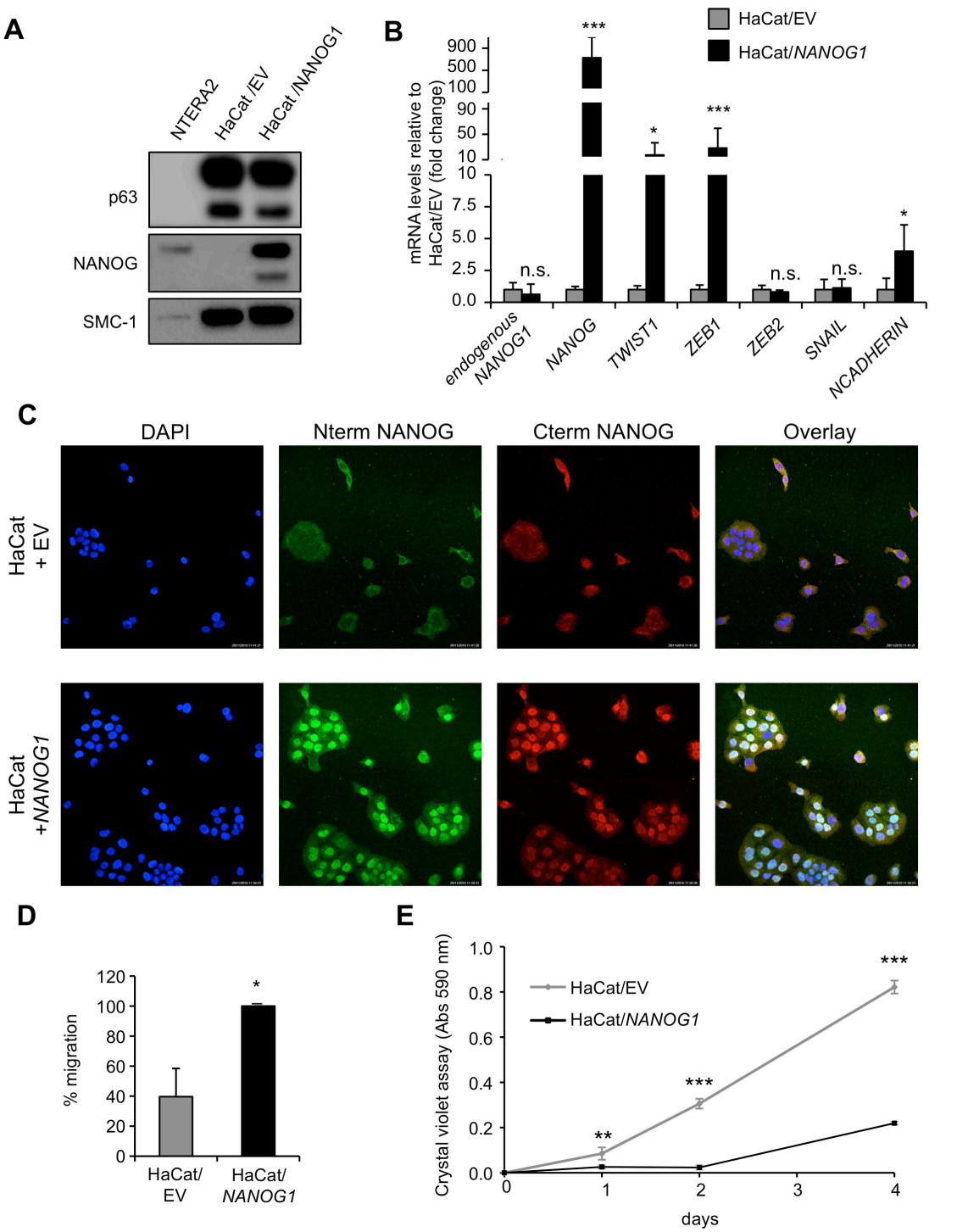


Figure 42. NANOG promotes EMT in epithelial cells

(A) Immunoblot of NANOG and p63 in nuclear extracts from HaCat cells infected with a human NANOG expressing vector (HaCat/*NANOG1*) or empty vector control (HaCat/EV). NTERA2 teratocarcinoma cells were used as a positive control. SMC-1 was used as a loading control.

(B) Relative mRNA levels of the indicated genes in HaCat cells treated as in (A). Samples were analyzed after puromycin selection. mRNA levels were normalized by *GAPDH* levels and then compared to levels HaCat/EV (shown as fold change). Values correspond to two independent biological replicates (n=2).

(C) Immunofluorescence of NANOG using two different antibodies (C-terminal and N-terminal detecting) of HaCat cells treated as in (A). DAPI was used to view the cell nuclei.

(D) Percentage of migration (24 hr) of HaCat cells infected with NANOG1 or empty vector (n=3). Areas were measured as percentage of migrated distance (measured across the scratch wound width).

(E) Growth-curve of HaCat cells overexpressing NANOG or empty vector. The experiment was repeated three independent times (each with a new NANOG infection of HaCat cells) obtaining similar results. The figure corresponds to one experiment. Quadruplicate wells were analyzed.

Values correspond to mean ± SD. Statistical significance was determined by the two-tailed Student's t test: n.s., non significant; (*) $P<0.05$; (**) $P<0.01$; (***) $P<0.001$.

Finally, we performed loss of function assays in a cellular model of esophageal squamous cell carcinoma (ESCC), given the fact that ESCCs are known to express high levels of endogenous NANOG (Bahl et al., 2012; Du et al., 2012; Fujiwara et al., 2013; Shimada et al., 2012; Yang et al., 2012; Zhang et al., 2012a). We silenced *NANOG* mRNA in TE2 ESCC cells using a *NANOG*-siRNA pool and compared it to its control (siRNA-scrambled). We confirmed *NANOG* knockdown by immunoblotting (**Figure 43A**) and by qRT-PCR **(Figure 43B)**. We also confirmed downregulation of some EMT transcription factors and stemness related genes such as *SOX2*, *LOXL1, ZEB1, ZEB2, WNT5A* and *CXCL12* **(Figure 43C)**.

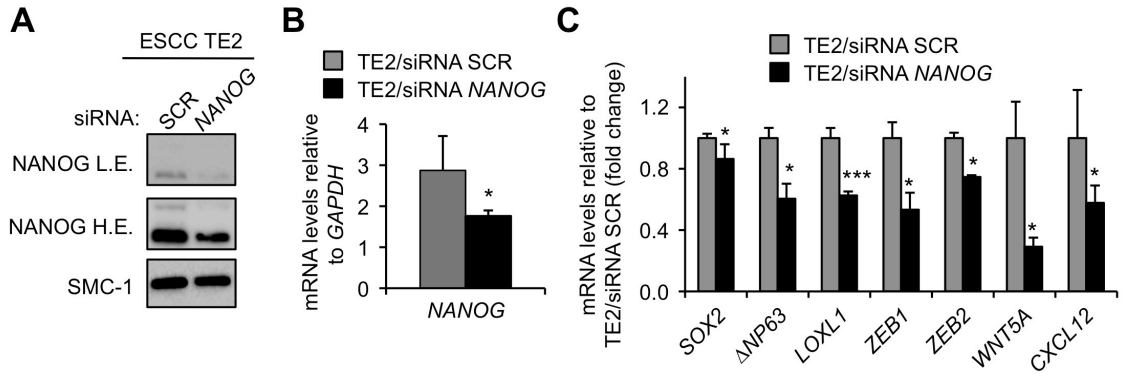


Figure 43. *NANOG* knockdown in ESCC TE2 cells leads to downregulation of EMT targets.

(A) Immunoblots of NANOG using TE2 (esophageal squamous cell carcinoma, ESCC, cell line) nuclear lysates cells transfected with pools of scrambled (siRNA SCR) or anti-NANOG (siRNA NANOG) siRNAs. Samples were analyzed 48 hr after transfection. SMC-1 was used as a loading control.

(B) Relative *NANOG1/P8* mRNA levels (abbreviated as *NANOG*) of TE2 cells as indicated in (A). mRNA levels were normalized by *GAPDH* levels. Values correspond to two transfections (n=2).

(C) Relative mRNA levels of the indicated genes in TE2 cells transfected as indicated in (A). mRNA levels were normalized by *GAPDH* levels and then compared to levels TE2/siRNA SCR (shown as fold change). Values correspond to two transfections (n=2).

Values correspond to mean ± SD. Statistical significance was determined by the two-tailed Student's t test: (*) $P<0.05$; (**) $P<0.01$; (***) $P<0.001$.

Together, our results unravel a novel function of NANOG in promoting squamous cell carcinoma conversion *in vivo* by directly activating EMT and stemness genes.

PART III
Role of NANOG variants in human cancer and reprogramming

3.1. Expression of NANOG paralogs in human cells

Research on *NANOG* is confounded by the presence of up to 11 known paralogs in the human genome (Booth and Holland, 2004). Many studies have characterized NANOG expression in different human tumors **(Table 1)**, but very few studies have distinguished which paralog is being expressed. The importance to differentiate among paralogs is that 3 of the NANOG variants, *NANOG1*, *NANOG2* and *NANOGP8*, give rise to functional protein and have aminoacid changes between them.

Our first approach was to design a strategy to identify the NANOG variants expressed in a specific sample. We designed a multi-*NANOG* RT-PCR based on two forward primers and one reverse primer with the potential to amplify most *NANOG* paralogs **(Figure 44A and 45)**. *NANOGP6* and *NANOGP7* were excluded due to their divergence compared to the rest of the variants. As a control of the PCR specificity we performed the PCR using genomic DNA extracted from human BJ fibroblasts. The PCR product was run on a gel, and the resulting PCR band was extracted, cloned into a pGEMT plasmid, transformed in competent DH5α *E. coli* and colonies were individually picked and sequenced **(Figure 44B)**.

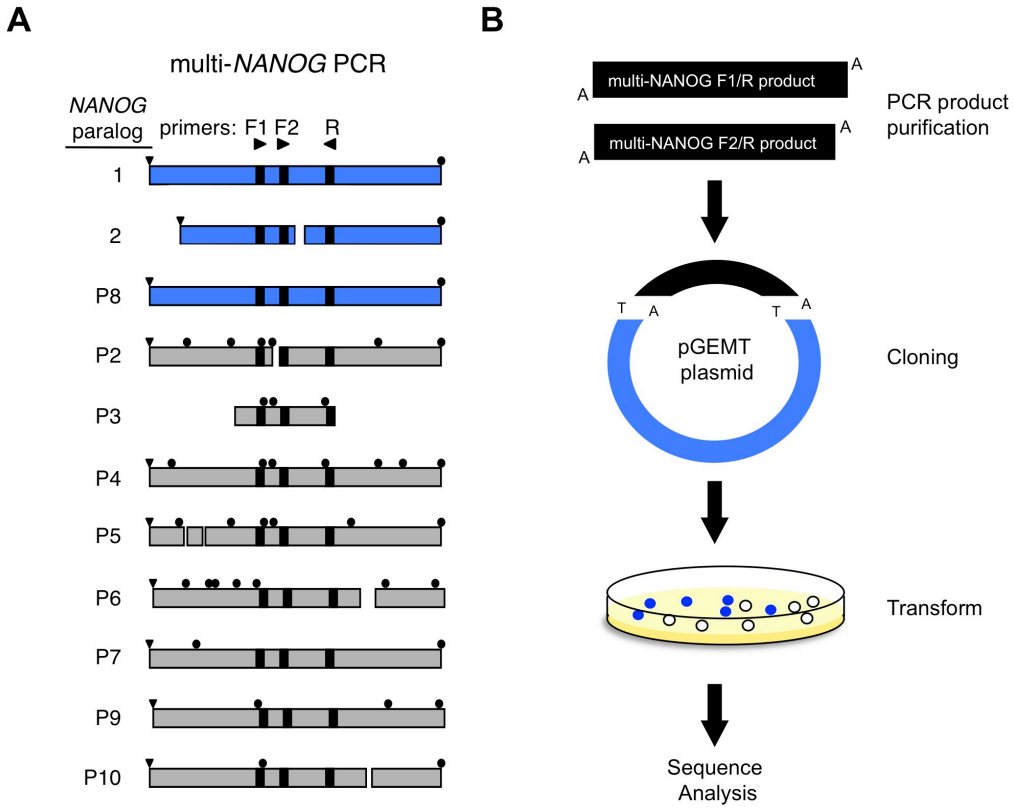
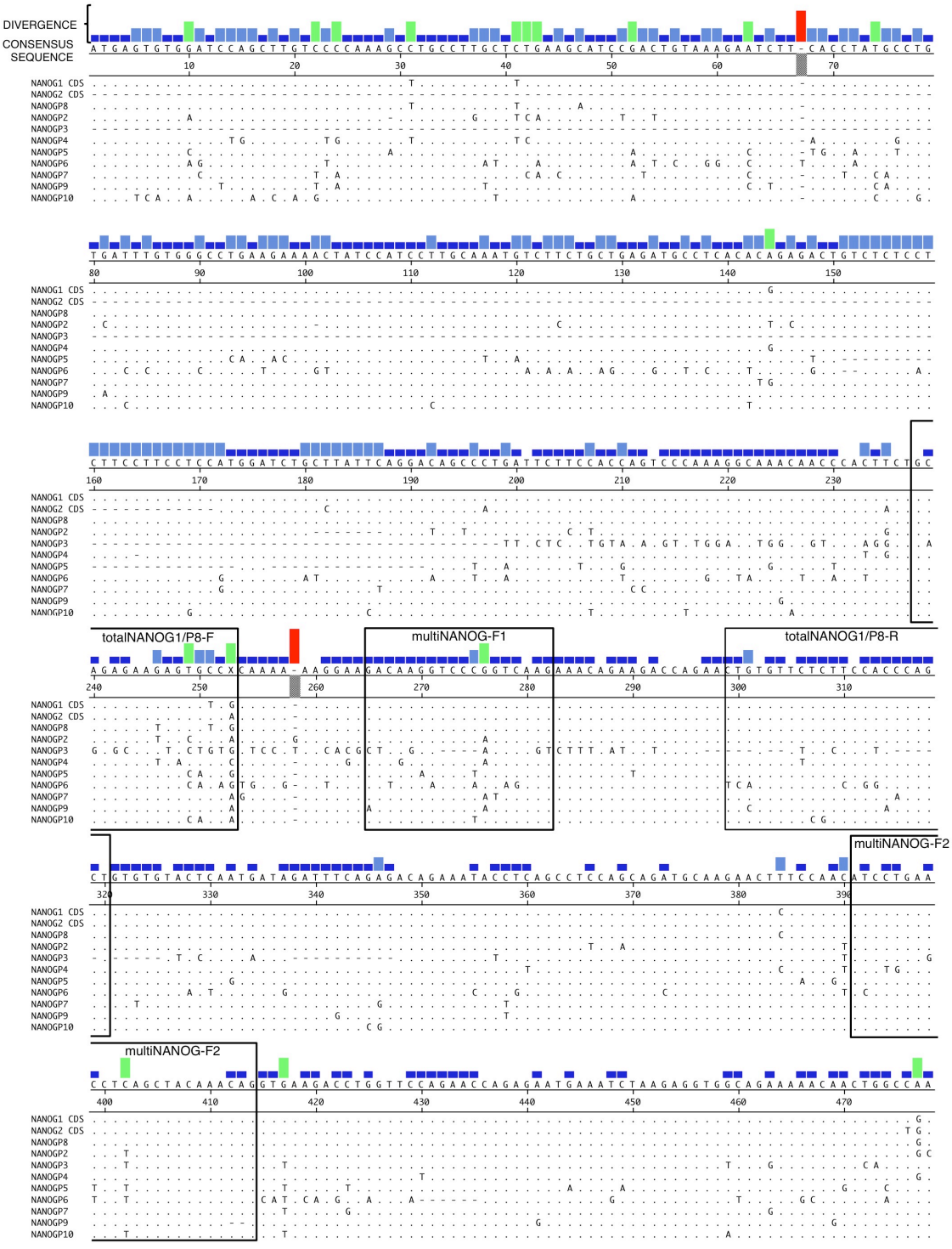


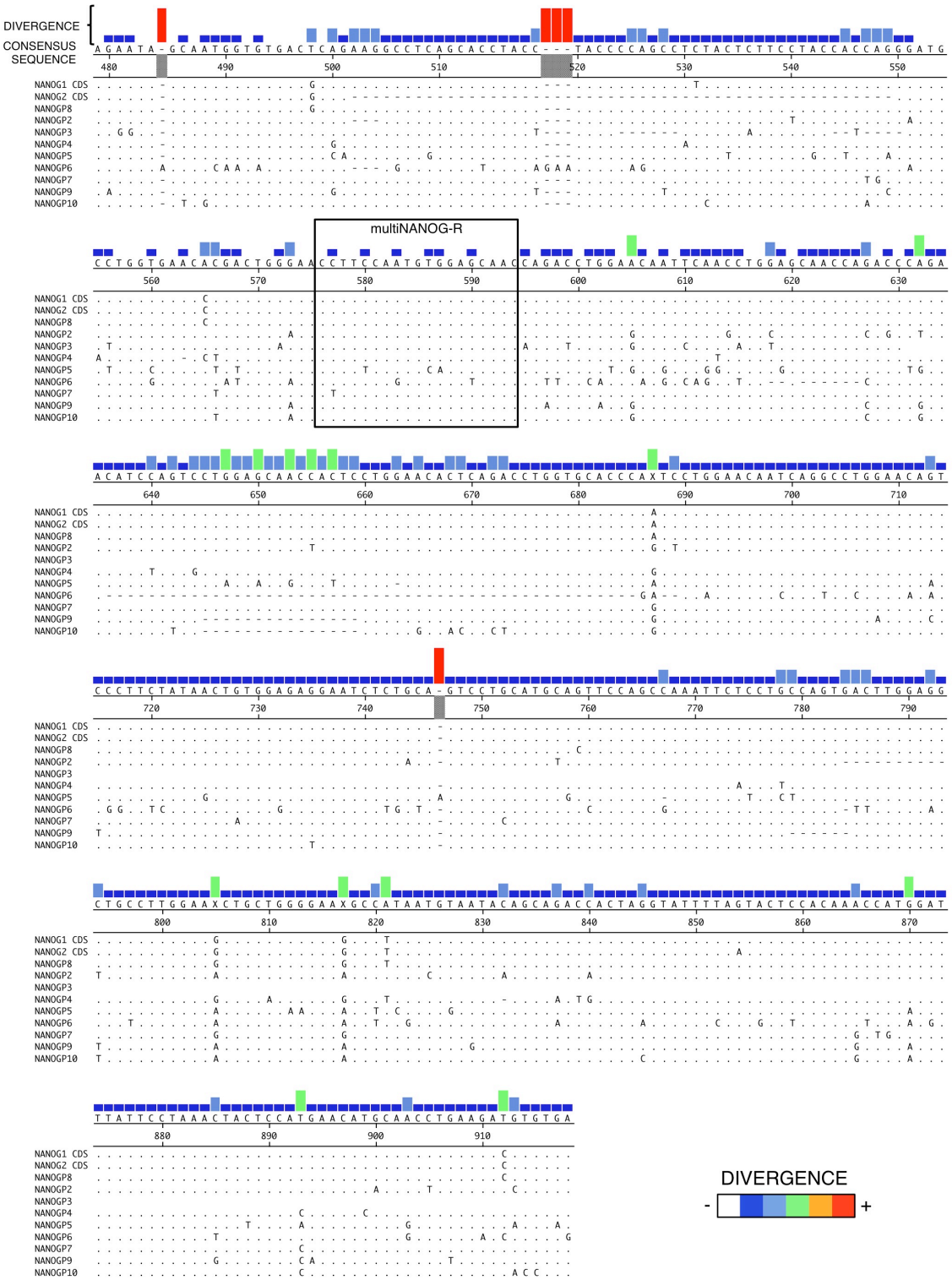
Figure 44. Multi-*NANOG* PCR

(A) Scheme of all *NANOG* paralogs and positions of the multi-*NANOG* primers (horizontal arrowheads and black boxes). The start codons are marked by vertical arrowheads, and stop codons by circles. *NANOG* variants with capacity to encode a functional NANOG protein are colored in blue. The remaining variants present numerous amino acid changes (not indicated), premature stop codons (circles) and gaps.

(B) *NANOG* variant identifying strategy. PCR product is cloned into pGEMT vector, transformed in *E. coli* and positive colonies are sent to sequence.



'.' symbol hides residues that match the consensus exactly.



'.' symbol hides residues that match the consensus exactly.

Figure 45. Alignment of all *NANOG* paralogs and primer binding sites
Alignment report comparing the coding sequence (CDS) of *NANOG1* to all *NANOG* paralogs. The upper sequence corresponds to the consensus sequence among all paralogs. The color and height of the bars above each nucleotide indicate the level of divergence between paralogs (see color key at the end of the alignment). The binding sites of the *NANOG* primers used in the paper are indicated with boxes. The symbol '.' hides residues that match the consensus sequence exactly. Gaps are indicated as '-'.

The analysis of 201 clones obtained from the genomic DNA of the BJ fibroblast cell line confirmed that the multi-NANOG PCR was able to amplify *NANOGP2, NANOGP3, NANOGP4, NANOGP5, NANOGP8* and *NANOGP10* of the *NANOG* variants. As expected, *NANOG1* and *NANOG2* were not amplified due to the presence of an intron between the primer-anchoring sites, but their presence was later on confirmed in other cDNA samples. Also, *NANOG1, NANOG2* and *NANOGP8* possess identical primer-anchoring sites and *NANOGP8* was correctly amplified. As an exception, *NANOGP9* was not amplified for unknown reasons. Therefore the multi-*NANOG* PCR strategy can detect and distinguish 8 out of the 11 *NANOG* variants, and most importantly all of the protein-coding variants **(Figure 46A)**.

We then decided to determine which *NANOG* paralogs were being expressed in a widespread panel of human cancer cell lines. HES2 (human embryonic stem cell line) and NTERA2 (human embryonic teratocarcinoma cell line) were used as a positive control since they are known to express *NANOG1* (Chambers et al., 2003; Mitsui et al., 2003) **(Figure 46A and 46B)**. It was important to ascertain that the cDNA was free of genomic DNA contamination since most of the paralogs are intronless. For this purpose, samples were treated with DNase I and additionally we performed the PCR on the original RNA samples (without addition of retrotranscriptase, -RT) **(Figure 46B)**. In agreement with previous studies (Ambady et al., 2010; Ishiguro et al., 2012; Jeter et al., 2009; Jeter et al., 2011; Uchino et al., 2012; Zbinden et al., 2010; Zhang et al., 2006a), intronless *NANOGP8* is the most abundant protein-coding *NANOG* in cancer cell lines (14 out of 17 cell lines analyzed). All cancer cell lines (17 out 17) expressed *NANOGP5* and many (7 out of 17) expressed also *NANOGP4*, an observation that again is in line with previous reports (Jeter et al., 2009; Zhang et al., 2006a) **(Figure 46A)**.

The *NANOG* expression pattern is different in pluripotent cells. In HES2 cells, *NANOG1* and *NANOG2* were the two variants identified. In NTERA2 cells, all protein-coding NANOG variants were detected: *NANOG1, NANOG2* and *NANOGP8* **(Figure 46A)**. This is the first study demonstrating *NANOG2* expression in embryonic stem cells. Of note, some cancer cell lines also showed expression of *NANOG1* in agreement with other studies (Ishiguro et al., 2012; Zbinden et al., 2010; Zhang et al., 2010).

From these analyses, we conclude that the majority of human cancer cells tested express *NANOGP8*.

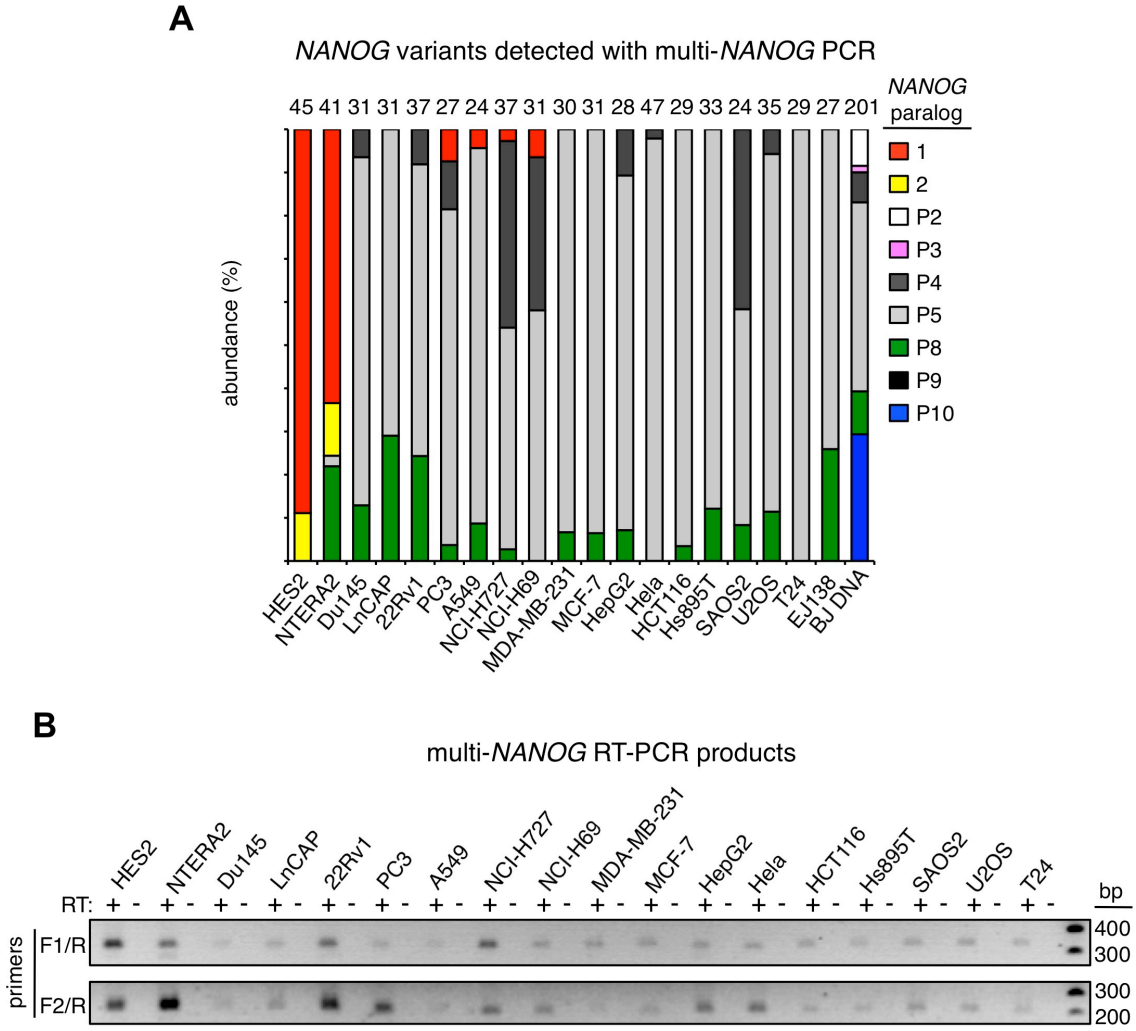


Figure 46. Analysis of NANOG variants in a panel of human cancer cell lines

(A) RT-PCR using multi-*NANOG* primers. cDNA was derived from RNA preparations of the indicated human cell lines. RNA samples not treated with reverse transcriptase (-RT) were used to monitor the absence of contamination by genomic DNA.

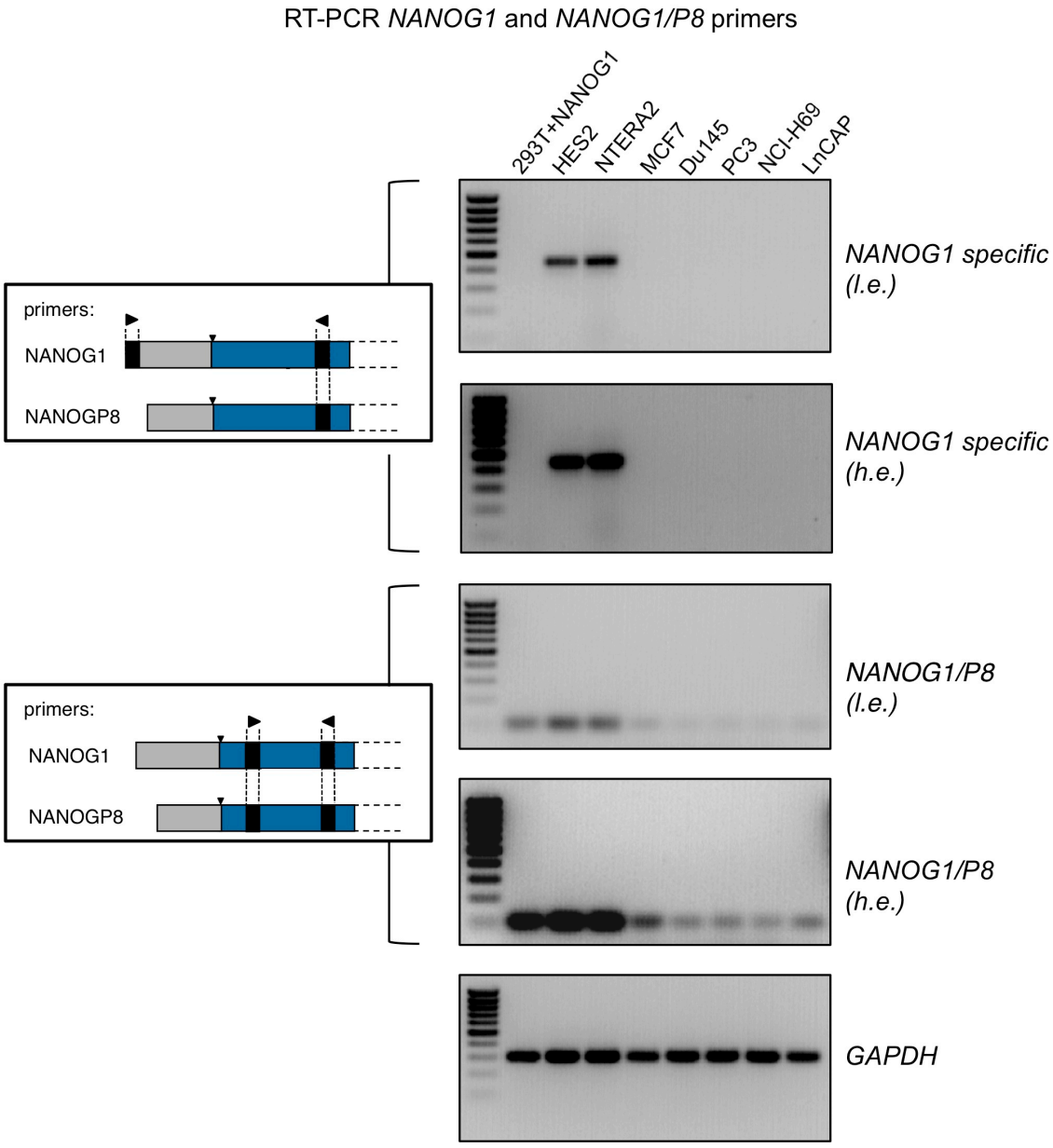
(B) Identification of *NANOG* variants after multi-*NANOG* RT-PCR and sequencing analysis of cDNA from human cell lines. Genomic DNA from BJ fibroblasts was used to evaluate the amplification potential of the different variants (BJ DNA). The number of sequences analyzed for each cell line is shown on the top and corresponds to the sum of the two PCR reactions F1/R and F2/R.

3.2. Expression of *NANOG1* and *NANOGP8* in human cells

The multi-NANOG RT-PCR strategy provides a qualitative rather than a quantitative estimate of the variants being expressed. The next approach was to obtain a quantitative analysis of the different protein-coding *NANOG* variants expressed in cancer.

NANOG1 and *NANOGP8* have very high homology, and only differ in 3 aminoacid changes at the protein level. This high similarity makes it difficult to design specific primers that can fully distinguish between both variants without interference from the remaining paralogs. We were able to design PCR primers that specifically amplify *NANOG1* thanks to a unique site present in the 5'UTR **(Figure 45 and 47)**. On the other hand we designed another pair of primers binding to the coding region that amplify simultaneously *NANOG1* and *NANOGP8* **(Figure 45 and 47)**.

Figure 47. Specificity of *NANOG1* and *NANOG1/P8* primers
Top: Semi-quantitative RT-PCR of *NANOG1* in indicated cell lines. The primers used selectively amplify *NANOG1* and their



position is indicated in the inset. The light gray box indicates the region of 5´UTR that is conserved between *NANOG1* and *NANOGP8*. The blue box indicates the coding region. Other symbols are as in Figure 44A. The forward primer binds to the 5'UTR of *NANOG1* at a region not conserved in *NANOGP8* 5´UTR. HES2 and NTERA2 cells were used as positive controls. 293T cells infected with a lentivirus expressing the CDS of *NANOG1* (without the 5'UTR) is also used as a control.
Bottom: Semi-quantitative RT-PCR of *NANOG1/P8* in indicated cell lines. The primers bind both to *NANOG1* and *NANOGP8* at the positions indicated in the inset. These primers do not amplify the other *NANOG* paralogs.
GAPDH is used as a loading control. l.e.: low exposition; h.e.: high exposition

PCR analyses show that *NANOG1* is abundantly expressed in pluripotent cells (HES2 and NTERA2) **(Figure 47 and 48A)**, but almost undetectable in most of the cancer cell lines. On the other hand, *NANOG1/P8* primers reveal detectable expression levels **(Figure 47 and 48B)**. We also ran the PCR product on a gel to corroborate that the products detected correspond to the correct size **(Figure 47)**. These data confirm that *NANOGP8* is expressed in human cancer cell lines.

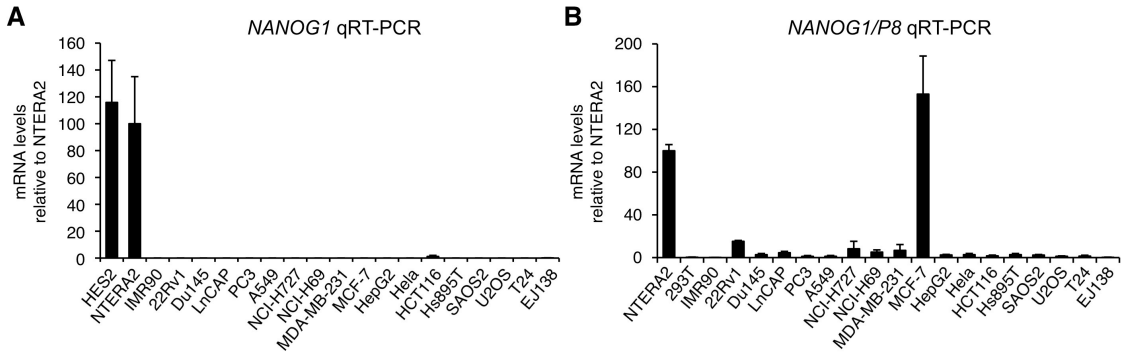


Figure 48. Expression of *NANOG1* and *NANOGP8* in human cell lines
(A) *NANOG1* expression levels in a panel of human cancer cell lines. HES2 and NTERA2 cells were used as positive controls.
(B) *NANOG1/P8* expression levels in human cancer cell lines.
mRNA levels were normalized by *GAPDH* levels and, then, compared to NTERA2 levels (100%). Values correspond to the average ± SD (n=6 technical replicates).

Based on the qRT-PCR data, we wanted to determine if *NANOGP8* expression gave rise to a detectable protein. We performed NANOG immunoblots on nuclear fractions from the previously studied cell lines. NTERA2 cells were used as a positive control. Several bands were observed in most cancer cell lines, a strong band below 50 kDa and a fainter double band at 37 kDa. This mobility pattern has already been reported and is in line with our results (Eberle et al., 2010; Jeter et al., 2009; Oh et al., 2005). Our previous results have shown that most cancer cell lines express *NANOGP8*, but not *NANOG1* (see above **Figures 46, 47 and 48**), therefore the ~37 kDa and ~50 kDa bands detected by immunoblot must derive from *NANOGP8*. Of note, the most abundant NANOG band in NTERA2 cells corresponds to the ~50 kDa band and is relatively abundant in cell lines 22Rv1, Du145, LnCAP, NCI-H727 and especially prominent in MCF-7 **(Figure 49)**.

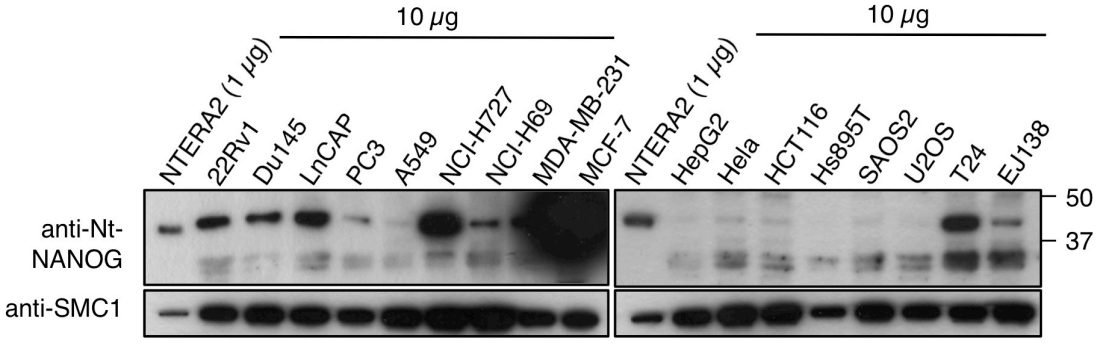


Figure 49. NANOG protein levels in human cancer cell lines
Nuclear extracts were immunoblotted using an antibody against the N-terminal region of NANOG (which only recognizes the proteins derived from *NANOG1* and *NANOGP8*, but not of *NANOG2*). Sizes of protein markers are indicated. NTERA2 cells were used as a positive control. For NTERA2 cells the amount of protein loaded was 1 µg; for all the other cell lines, the amount loaded was 10 µg. The cohesin subunit SMC1 was used as a loading control of the nuclear extracts.

To confirm that the ~37 kDa and ~50 kDa bands are produced by *NANOG* we performed knockdown of *NANOG* using three shRNAs targeting *NANOG* transcripts, with perfect matches towards *NANOG1* and *NANOGP8*. In the case of NTERA2 cells, each shRNA potently decreased the ~50 kDa band and, to a lower extent, also the ~37 kDa bands **(Figure 50)**. Additionally, we also performed immunoflurescence using the same antibody to assure specific

NANOG nuclear detection as well as correct silencing **(Figure 50)**. NANOG silencing in NTERA2 cell affected their morphology **(Figure 50)** in line with other reports (Greber et al., 2007).

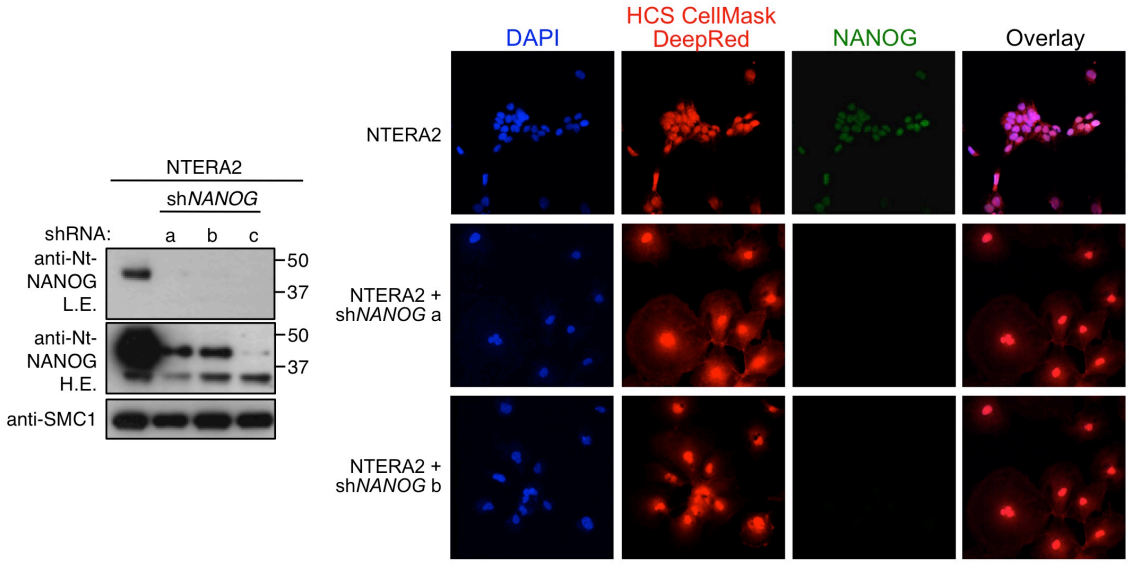


Figure 50. NANOG knockdown in NTERA2 cells
Left: Nuclear extracts from NTERA2 cells expressing different shRNA against *NANOG*, as well as scramble shRNA as a control (SCR), were immunoblotted using an antibody against the N-terminal region of NANOG. 5 µg were loaded for NTERA2 cells. Nuclear protein SMC1 was used as a loading control. H.E. and L.E. refer respectively to high and low exposition times after ECL incubation.
Right: Immunofluorescence of NANOG in NTERA2 cells after *NANOG* knockdown. Samples were analyzed after puromycin selection. DAPI was used to view the cell nuclei. HCS CellMask Stain was used to view the cell morphology (cytoplasm and nucleus).

A similar pattern was observed in cancer cell lines treated with a cocktail of the three shRNAs **(Figure 51)**. We performed qRT-PCR using the *NANOG1/P8* primers to detect silencing as well as immunoblots. Nuclear lysates were collected from cell lines 48 hours and 2 weeks after sh*NANOG* lentiviral infection. The 50 kDa band was more prominent and a clear decrease was observed after silencing. The ~37 kDa band is fainter and observed with higher exposition times but also decreases after *NANOG* knockdown in the case of cell line 22Rv1 **(Figure 51)**. Similar results were observed in HNSCC cells (see Part II, Figure 27A). These observations confirm that the ~37 kDa and ~50 kDa bands correspond to *NANOG*-derived proteins. At present, however, the molecular bases or significance of the various NANOG protein mobility forms remain speculative.

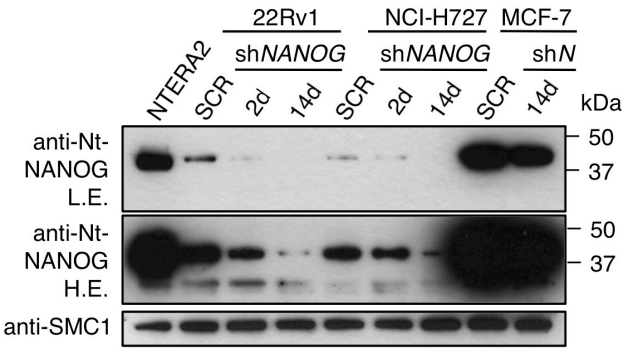


Figure 51. NANOG knockdown in cancer cell lines
Left: Nuclear extracts from cell lines 22Rv1, NCI-H727 and MCF-7 expressing shRNAs (pooled shRNAs a, b, and c) against NANOG, as well as scramble shRNA as a control (SCR). Two time-points were collected after sh*NANOG* infection of cell lines to view decrease of NANOG protein levels (2d and 14d refer to 2 days and 14 days post-infection of sh*NANOG*). 10 µg were loaded for cancer cell lines, and 5 µg were loaded for NTERA2 cells. Nuclear protein SMC1 was used as a loading control. H.E. and L.E. refer respectively to high and low exposition times after ECL incubation.

We can conclude from the mRNA and protein data that most of the human cancer cell lines analyzed contain detectable levels of NANOG protein derived from *NANOGP8*.

3.3. Reprogramming activity of *NANOGP8*

Once we have determined that *NANOGP8* is the main *NANOG* variant expressed in human cancer, our next approach was to view the pluripotency potential of *NANOGP8*. Several studies have already shown that *NANOGP8* contributes to the clonogenic and tumorigenic potential of human cancer cells (Jeter et al., 2009; Jeter et al., 2011; Uchino et al., 2012; Zbinden et al., 2010; Zhang et al., 2010). Therefore, we decided to focus on determining the reprogramming activity of *NANOGP8*. The NANOG protein encoded by *NANOGP8* presents two amino acid changes relative to the one derived from *NANOG1*, namely, at positions 16 and 253 **(Figure 52A)**. These two changes occur in functional protein domains of NANOG. In particular, residue 16 is part of a transcriptional repression domain, (Chang et al., 2009) whereas residue 253 is within a potent transcriptional activation domain (Do et al., 2009; Pan and Thomson, 2007). In addition, *NANOG1* presents a coding single nucleotide polymorphism (SNP) at position 82, which can encode for lysine or asparagine (Zbinden et al., 2010), and it is part of the transcriptional repression domain of NANOG1 (Chang et al., 2009). Of note, this residue is not polymorphic in *NANOGP8*, which always encodes an asparagine at position 82 **(Figure 52A)**. Site-directed mutagenesis was used to generate the two polymorphic versions of *NANOG1*, *i.e.* *NANOG1*^{K82} and *NANOG1*^{N82}. We also obtained a construct to express *NANOGP8*, as well as constructs carrying *NANOG1*^{K82} with each of the two *NANOGP8*'s changes, *i.e.* *NANOG1*$^{A16E;K82}$ and *NANOG1*$^{K82;Q253H}$. Expression from these vectors was confirmed by transfection of the retroviral vectors into 293T cells by immunoblotting and qRT-PCR **(Figure 52B and 52C)**. Surprisingly when we performed NANOG immunoblots from transfected 293T cells using two different antibodies, we observed clear mobility shifts between NANOG1, its point mutants and NANOGP8 **(Figure 52B). S**ubsequent analysis by qRT-PCR ensured that NANOG overexpression did not activate the endogenous locus **(Figure 52D)**.

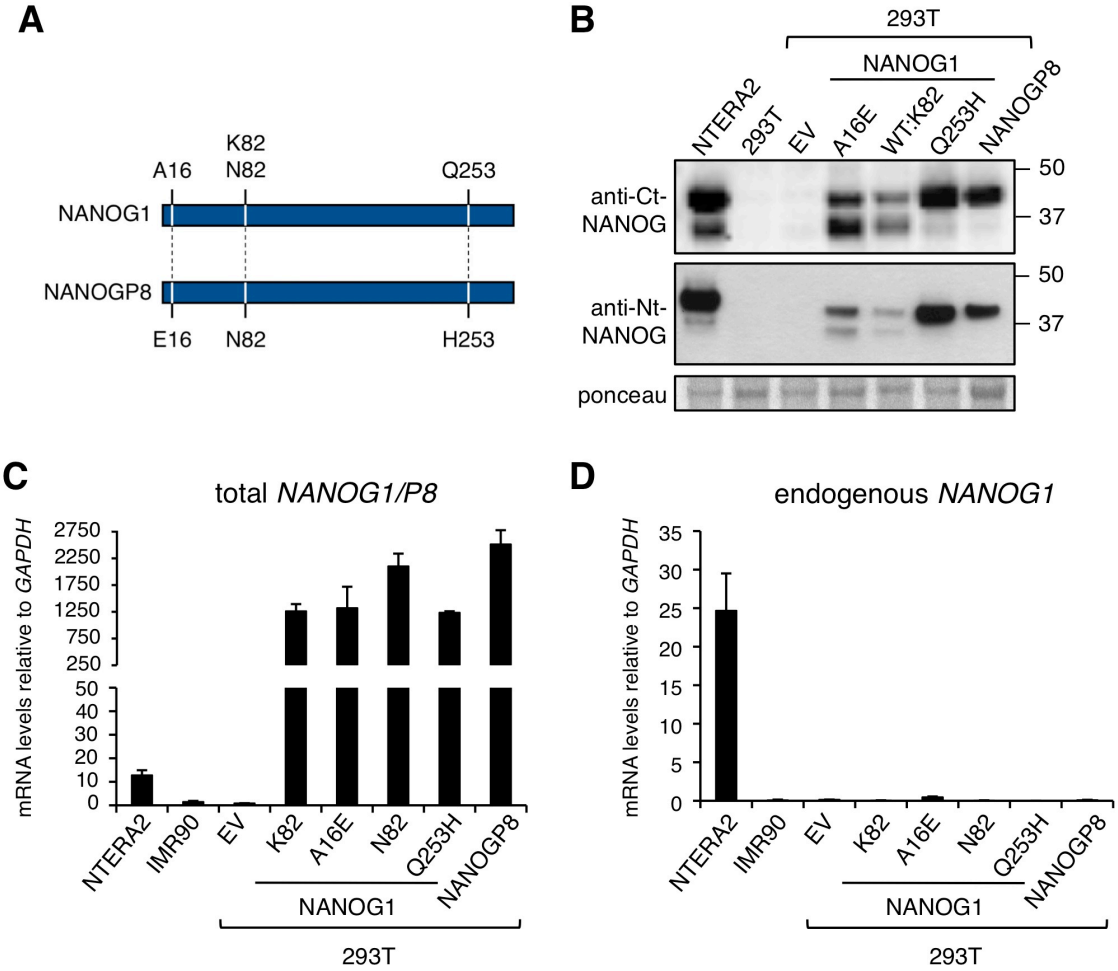


Figure 52. Expression of *NANOG* variants in 293T cells

(A) Cartoon showing differences between proteins encoded by *NANOG1* and *NANOGP8*. *NANOG1* presents two polymorphic variants at codon 82, one of which (N82) coincides with *NANOGP8*. *NANOGP8* has two amino acid changes relative to *NANOG1*^{N82} (E16 and H253).

(B) Immunoblots of NANOG using two different antibodies (C-terminal and N-terminal) in 293T cells transfected with control (EV: empty vector), *NANOG1*^{K82} and *NANOG1*^{N82} variants, *NANOG1*^{K82} point mutants A16E and Q253H, or *NANOGP8*.

(C) Expression levels of total *NANOG1/P8* in 293T cells transfected with the indicated retroviral constructs.

(D) Expression levels of endogenous *NANOG1* in 293T cells transfected with the indicated retroviral constructs.

(E) Teratocarcinoma cell line NTERA2 was used as a positive control and IMR90 fibroblasts as a negative control. Values correspond to the average ± SD (n=6 PCR replicates).

We began by testing the ability of the above-mentioned NANOG expressing constructs to improve the reprogramming efficiency of mouse cells. For this, we used MEFs derived from a reprogrammable transgenic mouse strain developed in our laboratory and carrying a polycistronic OSKM cassette under the control of the tetO element and the reverse transactivator (rtTA) in the *Rosa26* locus. Upon doxycycline addition to the medium these inducible-OSKM MEFs are efficiently reprogrammed (Abad et al., 2013). To establish a NANOG-dependent reprogramming assay in murine cells, we took advantage of the following two facts. On one hand, *Nanog* accelerates reprogramming (Hanna et al., 2009) and promotes the pre-iPSC to iPSC transition (Silva et al., 2009). On the other hand, only *bona fide* iPSCs are able to grow in the presence of MEK and GSK inhibitors (abbreviated as 2i), whereas 2i medium blocks the proliferation of pre-iPSCs **(Figure 53A)** (Silva et al., 2008; Ying et al.,

2008). To identify the conditions at which reprogramming is more dependent on NANOG, inducible-OSKM MEFs were infected with control (EV) or human *NANOG1*^{K82} and were subsequently treated with doxycycline (day 0). At different time points (day 3, 5, 7 or never), medium was supplemented with 2i and the final number of alkaline phosphatase positive (AP^{+}) colonies was scored at day 14. Addition of *NANOG1*^{K82} greatly increased the reprogramming efficiency and this was most apparent when adding 2i from day 3 **(Figure 53B)**. In particular, addition of 2i at day 3 almost completely prevented the emergence of colonies in the absence of *NANOG1*^{K82}. In contrast, the presence of *NANOG1*^{K82} allowed the formation of colonies in 2i medium confirming its ability to accelerate the emergence of fully reprogrammed iPSCs **(Figure 53B).**

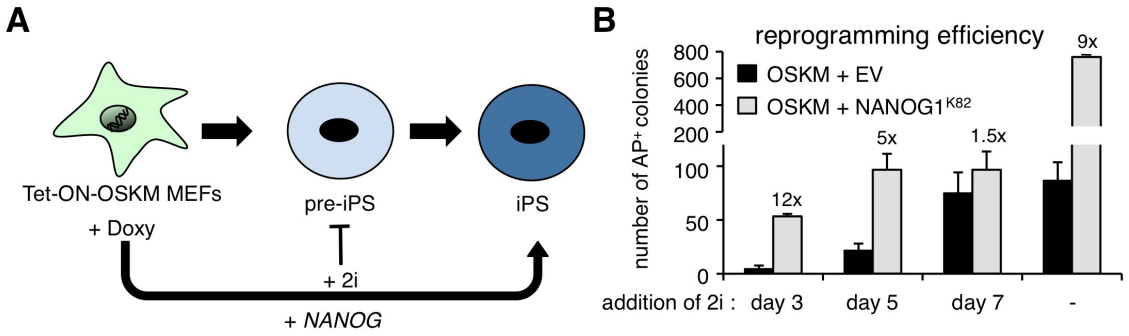


Figure 53. NANOG reprogramming assay

(A) NANOG is necessary for the second step during reprogramming of differentiated cells into fully reprogrammed iPSCs. Addition of 2i to the medium eliminates pre-iPS cells and promotes growth of naïve iPSCs.

(B) Quantification of the total number of alkaline phosphatase positive (AP^{+}) colonies. Inducible-OSKM MEFs were infected with either empty vector (OSKM + EV) or *NANOG1*^{K82} (OSKM + NANOG1^{K82}) and subsequently treated with doxycycline (day 0). Addition of 2i (CHIR99021 and PD0325901) was initiated at the indicated time or never (-) after doxycycline addition. AP staining was done at day 14. Three replicates were used. Values correspond to the average ± SD.

Having set up a NANOG reprogramming assay, inducible-OSKM MEFs were infected with retroviruses expressing mouse *Nanog*, human *NANOG1*^{K82}, *NANOG1*^{N82}, *NANOGP8*, as well as, *NANOG1*^{K82} point variants *NANOG1*$^{A16E;K82}$ and *NANOG1*$^{K82;Q253H}$. As mentioned above, we added 2i at day 3 and we scored AP^{+} colonies at day 14. The two *NANOG1* polymorphic variants were indistinguishable in this assay. Interestingly, *NANOGP8* had a reprogramming activity similar to *NANOG1*^{K82} and *NANOG1*^{N82} (**Fig. 54A and 54B**). In agreement with this, the two *NANOG1* point mutations, *NANOG1*$^{A16E;K82}$ and *NANOG1*$^{K82;Q253H}$, also showed similar reprogramming activity compared to the wild-type *NANOG1* variants **(Figure 54B)**. Finally, mouse and human *Nanog*/*NANOG* had a comparable reprogramming activity **(Figure 54B)**, extending the previous reported functional conservation of *Nanog*/*NANOG* in the pre-iPSC to iPSC transition (Theunissen et al., 2011). We confirmed similar retroviral expression levels in MEFs after infection **(Figure 54C)**. Of note, retroviral expression of *Nanog*/*NANOG* did not activate expression of the endogenous mouse *Nanog* gene **(Figure 52D)**.

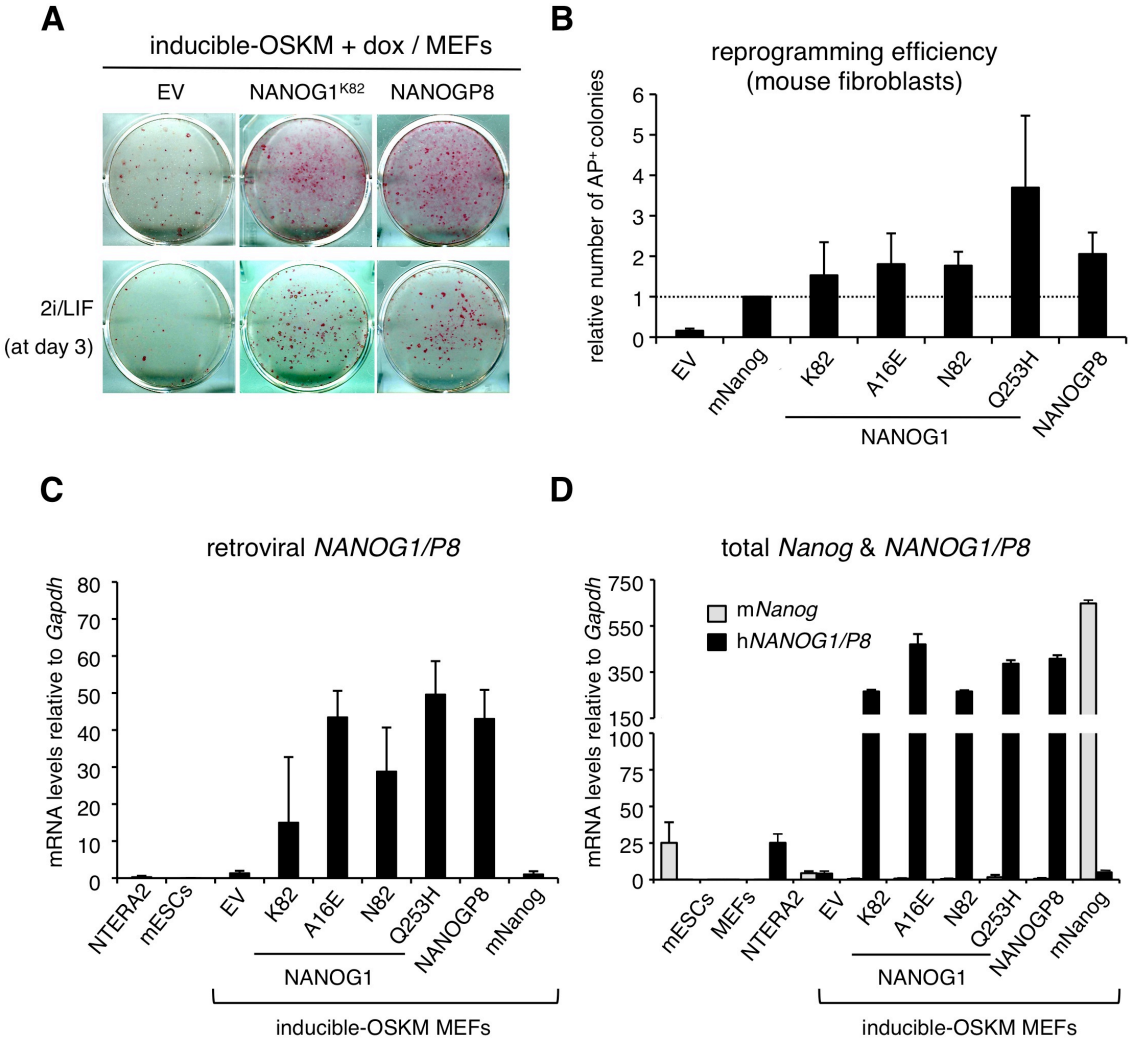


Figure 54. Activity of *NANOG* variants during reprogramming of MEFs

(A) Alkaline phosphatase staining of inducible-OSKM MEFs infected with empty vector (EV), *NANOG1*^{K82} or *NANOGP8* with subsequent doxycycline addition (day 0) cultured with or without 2i media (added from day 3).

(B) Relative reprogramming activity of inducible-OSKM MEFs infected with control (EV: empty vector), mouse *Nanog* (mNanog), *NANOG1*^{K82} and *NANOG1*^{N82} variants, *NANOG1*^{K82} point mutants A16E and Q253H, or *NANOGP8*. Reprogramming conditions were in the presence of 2i added from day 3 post-doxycycline. AP^{+} colonies were scored at day 14. Values correspond to the average ± SD. Four different biological replicates were performed (each using independent inducible-OSKM MEF isolates). For each replicate, the number of colonies was normalized to the number of colonies appearing in inducible-OSKM MEFs infected with mouse *Nanog*. Reprogramming efficiency was significantly increased by all the Nanog/NANOG constructs compared to the control (EV) (Student's t test $p<0.05$). The reprogramming efficiencies of the indicated NANOG constructs were not significantly different amongst themselves.

(C) Expression levels of retroviral *NANOG1/P8* in inducible reprogrammable OSKM MEFs infected with the indicated retroviral constructs. Values correspond to the average ± SD (n=4 full biological replicates each with an independent MEF isolate).

(D) Expression levels of total mouse *Nanog* and total human *NANOG1/P8* in inducible reprogrammable OSKM MEFs infected with the indicated retroviral constructs. Values correspond to the average ± SD (n=3 PCR replicates). NTERA2 and mouse embryonic stem cells (mESCs) were used as positive controls respectively, and MEFs were used as negative control.

We wanted to corroborate the above data in the reprogramming of human cells. For this, we used human primary foreskin fibroblasts and the three factor cocktail OCT4, SOX2 and KLF4 (OSK). Cells were retrovirally infected simultaneously with OSK retroviruses together with retroviruses carrying empty vector (EV), or the above-described *NANOG* constructs. Expression of NANOG variants was confirmed 48 hours post-infection **(Figures 55A and 55B)** and it did not activate expression of the endogenous *NANOG1* gene **(Figure 55C)**.

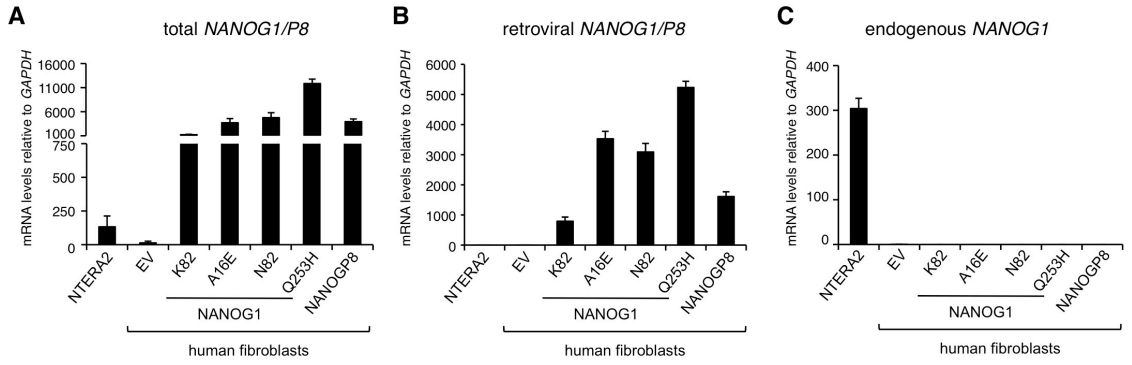


Figure 55. Expression levels of *NANOG* variants in retrovirally infected primary human fibroblasts

(A) Expression levels of total *NANOG1/P8* in primary human fibroblasts infected with the previously described retroviral *NANOG* constructs.

(B) Expression levels of retroviral *NANOG1/P8* in primary human fibroblasts infected with the previously described retroviral *NANOG* constructs.

(C) Expression levels of endogenous *NANOG1* in primary human fibroblasts infected with the previously described retroviral *NANOG* constructs.

NTERA2 was used as a positive control. Values correspond to the average ± SD (n=6 PCR replicates).

After 28 days, human iPSC colonies were clearly visible and had the standard morphology **(Figure 56A)**. Plates were stained with alkaline phosphatase and the number of AP^+ colonies was scored **(Figure 56B and 56C)**. Interestingly, the two polymorphic variants of *NANOG1* and *NANOGP8* increased reprogramming efficiency to a similar extent and, accordingly, the two *NANOG1* point mutants also had a similar effect **(Figure 56B and 56C)**.

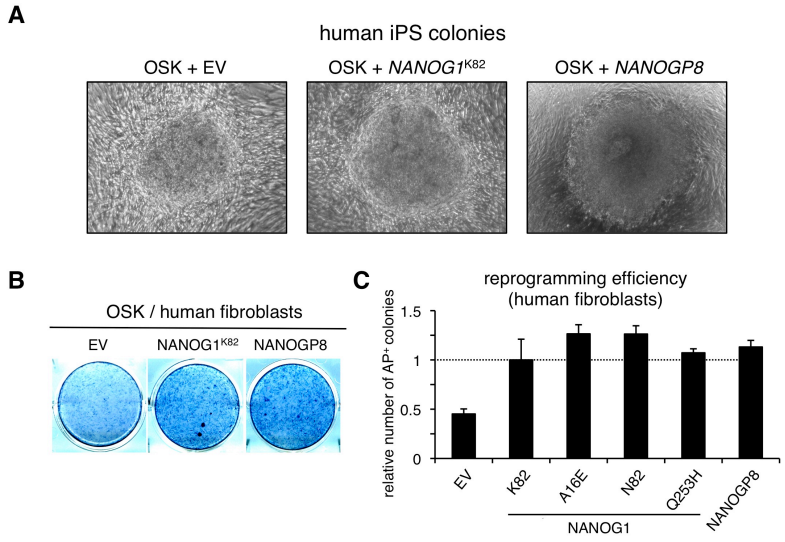


Figure 56. Activity of *NANOG* variants during reprogramming of MEFs

(A) Phase contrast microphotographs of representative human iPSC colonies at day 28.

(B) AP staining of human primary fibroblasts infected with OCT4, SOX2 and KLF4; as well as empty vector (EV); $NANOG1^{K82}$ or NANOGP8.

(C) Relative reprogramming activity of primary human fibroblasts infected with three factors OCT4, SOX2, KLF4 (OSK), together with control (EV: empty vector), with $NANOG1^{K82}$ and $NANOG1^{N82}$ variants, $NANOG1^{K82}$ point mutants A16E and Q253H, or *NANOGP8*. AP^+ colonies were scored at day 28. Values correspond to the average ± SD. Three replicates were performed and the number of colonies was normalized to the number of colonies appearing in fibroblasts infected with OSK and $NANOG1^{K82}$. The reprogramming efficiency of all the NANOG variants was significantly increased compared to the EV control (Student's t test $p<0.05$). The reprogramming efficiencies of the indicated NANOG constructs were not significantly different amongst themselves.

Together, these results demonstrate that *NANOGP8* possesses pluripotency capacity comparable to *NANOG1*.

DISCUSSION

Expression of NANOG has been always thought to be restricted to embryonic development and pluripotent cells (ESCs and iPSCs). In our work, we have uncovered new roles for NANOG in stratified epithelia and in cancer by use of an inducible *Nanog*-overexpressing mouse model. Additionally, we have analyzed a cancer-associated *NANOG* variant, *NANOGP8* and its role in pluripotency during reprogramming of differentiated cells.

1. NANOG and stratified epithelia

For a long time pluripotency factors were thought to only have a role during embryonic development. SOX2 perfectly exemplifies that a pluripotency transcription factor can have distinct roles in adult tissue homeostasis, depending on the tissue context and its binding partners. In our work we have discovered that NANOG is expressed in stratified epithelia of adult tissues. Specifically, it is expressed in the basal layer, which contains stem cell compartments that maintain the homeostasis of the epithelial tissue.

By using a mouse model that induces NANOG ubiquitous overexpression, we have observed that it can promote alterations in specific stratified epithelia such as forestomach and esophagus. Esophagus is unique compared to other epithelia such as skin, in the organization of its stem cell population. It contains a single layer of progenitor stem cells in the basal layer, which are in constant self-renewal to maintain the tissue (Doupe et al., 2012) (different from other adult stem cell populations which are normally quiescent (Blanpain and Fuchs, 2009)). Interestingly, NANOG overexpression is only effective in those tissues with endogenous levels of NANOG. This opens the possibility that tissue context is responsible for NANOG function in adult tissues.

NANOG-induced esophageal hyperplasia is also associated to an impaired differentiation, as shown by a decrease in differentiation markers and disorganization of the basal layer markers. This data indicates that NANOG could be promoting proliferation of the progenitor layer or dedifferentiation of the suprabasal layers. Already at short induction times (48 hr), microarray data revealed NANOG-overexpressing esophagus is capable of inducing a mitotic transcriptional program, in contrast to the liver.

1.1. NANOG and AURKA

Based on our microarray results, we have focused on the mitotic inducer Aurora kinase A (AURKA) as a potential mediator of NANOG's mitogenic activity. Consistent with this, we have proved that NANOG directly binds the *Aurka* promoter in ESCs and in keratinocytes. Moreover, we have demonstrated that NANOG can regulate AURKA expression levels and this

can directly affect proliferation of HNSCC cells. This is clearly shown by knockdown of NANOG in two HNSCC cell lines, which was associated to a significant reduction of AURKA protein levels and cells in S-phase.

AURKA and cell cycle

Aurora kinases play important roles during the mammalian cell cycle. AURKA is best known for its requirement in mitosis, where it is necessary for centrosome separation, mitotic entry and microtubule dynamics in order to build a bipolar spindle (Carmena and Earnshaw, 2003; Malumbres, 2011). Consistent with this, alterations of its expression lead to mitotic alterations that impair cell cycle. AURKA genetic ablation leads to early embryonic lethality because of its requirement in mitotic progression (Lu et al., 2008; Sasai et al., 2008). In the case of overexpression, AURKA can produce aberrant mitosis that lead to chromosomal aneuplodies (Bischoff et al., 1998; Meraldi et al., 2002; Zhou et al., 1998).

The function of AURKA is not only limited to mitosis, but it additionally has a role in DNA replication. This has been demonstrated by overexpression of AURKA which has been associated to increased proliferation in gastric cancer cells (Dar et al., 2009); consistent with a recent study showing that AURKA has a role in the induction of DNA replication by regulating the pre-replicative complex assembly (Tsunematsu et al., 2013).

Because of its critical role in mitosis and cell cycle, it has also been widely analyzed in cancer. *Aurka* has been identified as an oncogene in human cancer and it is amplified in a large number of cancer cell lines and tumors (Bischoff et al., 1998; Meraldi et al., 2002; Zhou et al., 1998). In human HNSCC this has also been observed, and overexpression of AURKA is associated with decreased survival, and a reduction of AURKA expression inhibits cell growth and increases apoptosis (Mehra et al., 2013). Its tumorigenic role in skin tumors *in vivo* has also been demonstrated, where ablation of *Aurka* in the basal layer of stratified epithelia leads to tumor arrest and decreased proliferation (Perez de Castro et al., 2013).

NANOG can upregulate AURKA, therefore it was important to determine if AURKA was responsible for the hyperplasia phenotype of the NANOG ubiquitous expressing mouse. Studies in our laboratory in collaboration with Ignacio Pérez de Castro and Marcos Malumbres (CNIO, Madrid) have shown that ubiquitous overexpression of AURKA in mice leads to increased proliferation and aneuploidy of the basal layer of the esophagus (D.P., *in revision*). These results show AURKA is a potential mediator for the mitotic function of NANOG. Of note, AURKA overexpression showed a milder phenotype (proliferation and aneuploidy) than the one observed for NANOG overexpression, suggesting that although AURKA is an important effector, it is not the only downstream target of NANOG.

AURKA and polarity

Aurora kinases have also been associated to regulation of the balance between symmetric and asymmetric cell divisions, because of their role in the establishment of polarity during cell divisions (Malumbres, 2011). Asymmetric and symmetric divisions are important for the maintenance of stem cell populations and tissue homeostasis in adults. *In vivo*, this is determined by the stem cell niches that possess polarized cells regulating the orientation of cell division to determine the fate of daughter cells (Neumuller and Knoblich, 2009).

The epithelium is a good example of polarized tissues. Here the cells of the basal layer are polarized establishing two membrane domains, the apical and basolateral membranes **(Figure 57)**. In symmetric divisions the division plane is parallel to the polarity axis (basement membrane), and cell fate constituents will be equally segregated into daughter cells remaining in the basal layer (stem cell niche). On the other hand, asymmetric divisions consist on orientating the spindle perpendicular to the basement membrane, where daughter cells inherit different contents and diverge in differentiation **(Figure 57)** (Blanpain and Fuchs, 2009; Neumuller and Knoblich, 2009; Niessen et al., 2012). AURKA is involved in asymmetric divisions due to its role in centrosomal and spindle function, and by regulating localization of asymmetric marker Numb (Wirtz-Peitz et al., 2008). AURKA hypomorphic loss in neuroblasts of flies impairs asymmetric localization and increases symmetric cell division (Berdnik and Knoblich, 2002; Wang et al., 2006a).

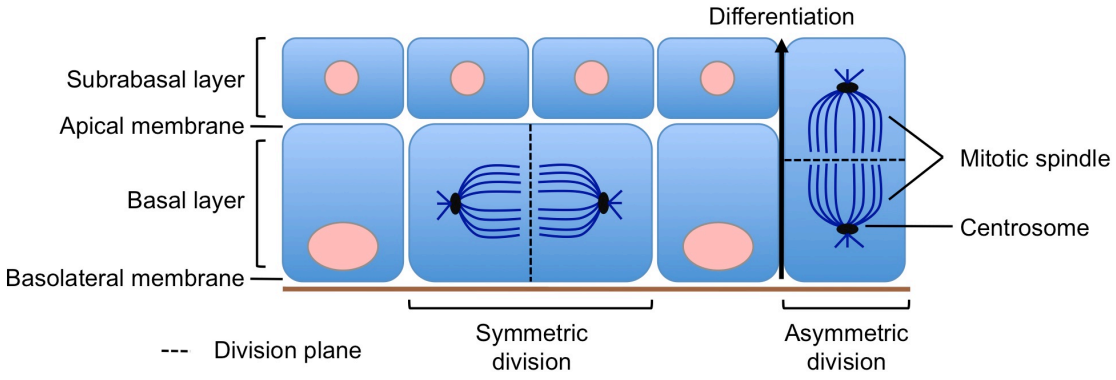


Figure 57. Symmetric and asymmetric divisions in epithelia.
During epithelial homeostasis, asymmetric cell divisions occur with the plane of division parallel to the basement membrane and gives rise to two different daughter cells, one remains a stem cell, whereas the other becomes committed to terminal differentiation. To maintain the pool of stem cells, symmetric cell division occurs resulting in two equal daughters, also known as stem cell self-renewal.

In mammals, a recent study has also shown that AURKA is essential to regulate stem cell fate in mouse mammary epithelium, through the function of AURKA in mitotic spindle regulation (Regan et al., 2013). All this evidence, points to AURKA as a key player in determining cell fate by controlling symmetric and asymmetric division, and maintenance of the stem cell population.

An imbalance of these divisions has also been related to cancer. An increase of symmetrical divisions promotes carcinogenesis in *Drosphila* (Castellanos et al., 2008; Caussinus and Gonzalez, 2005), although it remains as a hypothesis to be confirmed in mammals. Since we have demonstrated NANOG can regulate AURKA in epithelial cells, this opens the possibility that it could also have an effect in tissue homeostasis, maintenance of the stem cell niche and tumorigenesis thorough AURKA activity as a "polarity related kinase". This is an interesting line of work that should be further studied, for which a conditional knockout model for NANOG would be required.

All together, our observation of direct regulation of NANOG over AURKA opens the possibility that NANOG can have a causal effect in tumorigenesis. AURKA inhibitors are potential drugs that are now used in clinical trials in cancer, including HNSCC (Manfredi et al., 2011; Mehra et al., 2013). When we knockdown NANOG in HNSCC cells, proliferation decreased and cell cycle is arrested together with significant reduction of AURKA protein levels. Further studies in our laboratory in collaboration with the group of Juana Maria García-Pedrero, Marta Sánchez-Carbayo and Manuel Rodrigo-Justo have found a positive correlation between NANOG and AURKA expression in HNSCC and ESCC. These results show that NANOG could be used as an informative biomarker in cancer to predict the response to AURKA inhibitors.

1.2. NANOG long-term overexpression in stratified epithelia

Regarding long-term overexpression of NANOG in mice, we have found different results depending on the compartments where we express NANOG: (1) In the ubiquitous overexpression model (ubiquitous-*Nanog*), mice die within weeks and show clear hyperplasia of stratified epithelia as well as thymus atrophy. (2) In the epithelial basal layer overexpression model (K5-*Nanog*), there is no apparent phenotype even after 9 months of doxycycline administration. This should be analyzed in detail, but several clues could account for this:

(1) The pattern of expression is different: In the ubiquitous-*Nanog* mice, NANOG is not only expressed in the basal layer but also in the suprabasal layers. In the K5-*Nanog* mice, NANOG is only expressed in the basal layer.

(2) The hematopoietic stem cell phenotype: In the ubiquitous-Nanog mice show atrophy of the thymus. In the K5-Nanog, the thymus is normal and presents no phenotype. A previous report has shown that *Nanog* is expressed at low levels in some hematopoietic progenitor lineages (B-cell lineage; sub-population of the myeloid lineage, and thymus double positive stage cells). In this same report, transplant assays done with *Nanog* overexpressing hematopoietic stem cells

(HSC) show that NANOG inhibits T-cell development and promotes dominance of γδT-cells in the thymus leading to thymus atrophy. Furthermore, long-term overexpression of NANOG in HSCs leads to development of a lymphoproliferative disorder (Tanaka et al., 2007). This phenotype could be key for the epithelial phenotype observed in the *Rosa26::rtTA, Col1a1::tetO-Nanog* mice. γδT-cells constitute a small percentage of the T cell population but are enriched (> 50%) in epithelial cell-rich compartments like the skin, digestive tract and reproductive organ mucosa. These T-cells have been implicated as early and rapid responders to tissue damage (Havran and Jameson, 2010; Komori et al., 2006). They are in close contact to epithelial keratinocytes, and upon activation they can secrete cytokines and chemokines which play roles in inflammation, and recruit immune cells to the tissue (Macleod and Havran, 2011). Based on these results, we can hypothesize that in our ubiquitous mouse model, NANOG overexpression also forces HSCs to a γδT-cell fate. Moreover, we also observe thymic atrophy in our mice. A shift towards γδT-cells can also be driving a hyperplasic effect in epithelia modulated by inflammation. In our case, the K5-inducible *Nanog* mouse model only showed a distinct phenotype upon TPA application, known to promote proliferation and inflammation. In this case we also see an increased hyperplasia and impaired differentiation of the basal layer in the skin.

Interestingly, another group also tried to answer the role of *Nanog* in stratified epithelia and cancer by generating a transgenic mouse model overexpressing the human *NANOG* variant, *NANOGP8* under the control of the cytokeratin 14 (K14) promoter. These K14-*NANOGP8* transgenic mice show partial perinatal lethality. However K14-*NANOGP8* mice that survive grow normally as adults showing subtle abnormalities such as atrophy in the skin and tongue hyperplasia. In agreement with our results, they do not develop spontaneous tumors (Badeaux et al., 2013).

1.3. NANOG partners in stratified epithelia

Another question is why NANOG overexpression has a specific effect in stratified epithelial tissues. As already mentioned, SOX2 binding to different partners regulates its roles in adult tissues. Of particular interest, P63 has been shown to bind to SOX2 in SCC cells and drive a different transcriptional program compared to ESCs (Watanabe et al., 2014a). Our results reveal that NANOG follows the same expression pattern as P63. In addition, we have also seen that NANOG interacts with ΔNp63α in ESCC cells. We will further address this issue by analyzing ChIP-seq of NANOG in ESCC cells, and clearly identifying the NANOG transcriptional program in SCC.

1.4. NANOG and tissue homeostasis

Our approach to examine the role of *Nanog* in tissue homeostasis has consisted on using an inducible-*Nanog* overexpression mouse model. Forced overexpression of ubiquitous *Nanog* results in hyperplasia of squamous epithelia, such as esophagus and forestomach. With these results we confirm that *Nanog* can have an effect in epithelia, but what still remains to be elucidated is the role and requirement of endogenous *Nanog* for tissue homeostasis. This can only be addressed with a loss-of-function model. As *Nanog* deletion is early embryonic lethal, a conditional KO strategy should be used. In our laboratory, this mouse model is currently being developed to answer this question.

In summary we have observed that NANOG is expressed in adult tissues in stratified epithelia, and ubiquitous overexpression of *Nanog* causes hyperplasia and induction of a mitotic transcriptional program specifically in stratified epithelia (most accentuated in the mucosal epithelia). The specificity of NANOG in stratified epithelia suggests that it might bind to partners that regulate its function, being ΔNp63α and SOX2 good candidates. Together, these results indicate that NANOG has a lineage-restricted mitogenic function in stratified epithelia.

2. NANOG and cancer

As mentioned previously, NANOG can directly regulate AURKA expression, which could have a role in tumor initiation. Our results show that *Nanog* expression alone is not sufficient to induce tumors. Other well-known oncogenes have also been demonstrated to not spontaneously give rise to tumors when overexpressed. This is the case for *cMyc*, not only a well-established oncogene, but also one of the reprogramming "Yamanaka" transcription factors. *cMyc* overexpression leads to increased epidermal proliferation and ultimately to depletion of stem cells, rather than expansion to form tumors (Arnold and Watt, 2001; Gandarillas and Watt, 1997; Waikel et al., 2001). K14-*NANOGP8* overexpressing transgenic mice phenocopy this, where NANOG overexpression during a long period of time shows a depletion of the stem cells and lack of tumor formation (Badeaux et al., 2013).

Hence, we decided to use a chemical carcinogenesis approach to view NANOG's role in tumorigenesis. We demonstrate that NANOG can promote progression in a resistant mouse background upon chemical carcinogenesis by activating an EMT and stemness program. These results unravel NANOG as an oncogenic driver in skin tumors and strengthens the idea of pluripotent transcription factors as tumorigenic in a cancer context (Suva et al., 2013). *Nanog* overexpression alone is not sufficient to induce alterations in skin. But in *Nanog*-overexpressing papillomas, we could already determine some histological differences including disorganization

of the basal layer and atypia, as shown by epithelial stem cell marker P63 IHC. This suggests NANOG requires additional events (*Ras* mutation and TPA mitogenic stimulation) to act as a driver of oncogenesis.

2.1. NANOG, EMT and CSCs

Cancer cells and pluripotent cells share similarities. Both reprogramming to induced pluripotency and malignant transformation can be defined as stepwise processes, where only a small percentage of cells undergo the necessary changes to create either iPSC or cancer cells. iPSCs and cancer cells share properties such as unlimited proliferation, self-renewal and tumorigenic potential (Apostolou and Hochedlinger, 2013). Although still controversial, there are several lines of evidence showing the idea that tumor heterogeneity relies on a fraction of cells that has the capacity to self-renew and give rise to the different lineages of cancer cells known as cancer stem cells (CSC). Cell plasticity refers to the ability of cells to reversibly change their phenotype, in tumor cells this could refer to a certain subpopulation of cells which dedifferentiates and acquires CSC properties under certain conditions (Friedmann-Morvinski and Verma, 2014). Related to cancer, tumor plasticity and CSCs have also been related to EMT and to the invasion-metastasis cascade (Mani et al., 2008). EMT and its reverse process mesenchymal-epithelial transition (MET) are common processes both in development and in cancer. In development, cells can migrate after undergoing EMT and once at their destination undergo MET to give rise to different tissues. Likewise, this process also occurs in cancer, where tumor cells undergo EMT, extravasate into lymphatic and blood vessels, and subsequently extravasate following MET to form metastasis (Nieto, 2013). These steps required for metastasis formation require certain plasticity from cancer cells, since EMT TFs must be downregulated to form the metastasis, and once at the secondary site return to an epithelial cell resembling the primary tumor.

Our results show that *Nanog*-overexpressing papillomas have an increased expression of EMT TFs. There are already several reports linking NANOG to an EMT phenotype in cancer cells **(Table 1)**, including one study showing that NANOG can directly regulate TGF-ß1 expression in melanoma cells, a potent inducer of EMT (Hasmim et al., 2013). Our study is the first to show that *Nanog* overexpression is capable of driving an EMT program in skin tumors. Our RNA-seq data revealed that in papillomas, EMT inducer *Prrx1* (Ocana et al., 2012) was upregulated. Further analysis by qRT-PCR of *Nanog*-overexpressing papillomas and carcinomas confirmed upregulation of other EMT drivers such as *Zeb1, Zeb2* and *Twist1*, as well as downregulation of *E-cadherin*, a key property of EMT (Sanchez-Tillo et al., 2012).

PRRX1

PRRX1 was recently discovered as a driver of EMT in breast cancer, and its downregulation is essential for extravasated cells to colonize distant tissues during metastasis (Ocana et al., 2012). In relationship to the finding of *Prrx1* as an upregulated target in our Nanog-overexpressing papillomas, we did not find *Snail* or *Slug* upregulation in *Nanog*-overexpressing papillomas, carcinomas or keratinocytes. *Snail* and *Slug* are master regulators of the EMT process (Cano et al., 2000). However, a recent study has hypothesized that *Prrx1* can induce a different type of EMT independent of *Snail* and *Slug*. These two types of EMT (Snail-type and Prrx1-type) could be either parallel or exclusive processes in a tumor (Ocana et al., 2012). There are some theories linking EMT with stemness properties, where dedifferentiated cells in a tumor are the ones giving rise to cells with migration capacity (Brabletz et al., 2005). New studies show metastasis requires both processes (EMT and MET) (Ocana et al., 2012; Tsai et al., 2012). This suggests that metastatic cells require cell plasticity, a property that could be given by stemness factors. In our case, we have performed RNA-seq of the whole tumor population and cannot distinguish different subpopulations. It would be interesting if we could capture circulating tumor cells in the bloodstream of *Nanog*-overexpressing mice to confirm the EMT and dissemination hypothesis. Likewise, it would be appealing to analyze if there are tumor subpopulations which are more dedifferentiated to understand mechanistically what *Nanog* is promoting in a tumor: EMT coupled to stemness; or two subpopulations, one linked to EMT (*Prrx1*-related), and another to stemness.

ZEB1, ZEB2 and TWIST1

Our results also show that *Nanog*-overexpressing papillomas and carcinomas regulate other key EMT TFs, including *Zeb1, Zeb2* and *Twist1*. ZEB1 and ZEB2 form part of the ZEB family that are zinc finger/homeodomain proteins. They are known to mediate EMT by repression of epithelial markers (either directly or through regulation of miRNAs) and activation of mesenchymal markers (Thiery et al., 2009). TWIST1 is also a basic helix-loop-helix with DNA binding capacity protein with a “Twist box” at its C-terminal, which can regulate gene expression either by activation or repression of target genes involved in EMT (Thiery et al., 2009). Our results have shown that NANOG can upregulate the ZEB factors (*Zeb1* and *Zeb2*) and *Twist1* in a cell-autonomous manner. Interestingly, we have observed NANOG can bind to the promoters of *Zeb2* and *Twist1* providing a direct link of NANOG with EMT induction.

2.2. NANOG and stemness

Our RNA-seq data showed *Nanog*-overexpressing papillomas are enriched in genes related to stemness. Interestingly, two well-known markers related to CSCs came up, *ALDH1* and *CXCL12/SDF-1*. Among our targets we also observed upregulation of stemness related genes *LGR5, PDGFRA* and *PITX2*.

ALDH1

ALDH1 has been well characterized to identify breast stem cells and breast CSCs associated to malignancy (Ginestier et al., 2007). It has also been identified to mark CSC populations in melanoma, liver, lung, pancreatic and prostate tumors (Medema, 2013). In HNSCC, reports have demonstrated ALDH1 expression had a positive correlation to staging and a negative correlation to patient outcome (Chen et al., 2009). ALDH1 is a member of the aldehyde dehydrogenase family, with functions ranging from detoxification by oxidation of intracellular aldehydes (Magni et al., 1996) to early stem cell differentiation by conversion of retinol to retinoic acid (Chute et al., 2006). Related to the above association of stemness and EMT, ALDH1 has also been connected to EMT, where $ALDH1^+$ HNSCC CSCs were found to increase expression of EMT inducer *Snail* (Chen et al., 2009).

CXCL12

Regarding CXCL12 (also known as stromal cell-derived factor-1, SDF-1), chemokines have long been known to determine the development of primary tumors and metastasis (Balkwill, 2004). Chemokine receptors are able to induce directional migration of cells toward chemokines (chemotactic cytokines), and were initially identified for their ability to cause directed leukocyte migration to sites of inflammation (Baggiolini, 1998). Further studies revealed their importance in other processes such as embryogenesis, angiogenesis, hematopoiesis, HIV1-infection and cancer (Balkwill, 2004; Doitsidou et al., 2002; Feng et al., 1996; Romagnani et al., 2004). CXCL12 is a homeostatic chemokine, which specifically binds and initiates signaling through receptor CXCR4, a well-known 7 transmembrane domain G-protein-coupled receptor (GPCR). The CXCR4/CXCL12 pathway has shown to play a critical role in cancer progression and metastasis (Teicher and Fricker, 2010). As mentioned before, CXCL12 has been associated to stemness, since its receptor has been identified as a CSC marker in pancreatic cancer, where $CD133^+$ $CXCR4^+$ cells were found to be essential in determining tumor metastasis (Hermann et al., 2007). CXCR4 is widely expressed in many tumor types and is responsible for metastasis to the lung, liver and bone marrow (Zlotnik et al., 2011). CXCL12 can also stimulate survival and growth of neoplastic cells in a paracrine fashion and promote tumor angiogenesis by attracting endothelial cells (Burger and Kipps, 2006). The relationship of NANOG to this pathway is also

related to its role in embryogenesis. The CXCL2/CXCR4 pathway is key for PGC migration (Doitsidou et al., 2002), where *Nanog* is also known to be essential. In agreement with this, one report found that NANOG could directly regulate the expression of *Cxcr4* in PGCs and mediate their migration in medaka (Sanchez-Sanchez et al., 2010). Our results indicate NANOG can directly regulate CXCL12 expression, which can play a driving effect in progression of squamous tumors. The CXCL12/CXCR4 pathway is a current target for anticancer therapy since it is known to activate oncogenic signaling pathways PI3K/Akt, IP3 and MAPK (Duda et al., 2011); if NANOG is upstream of this pathway it could also be a potential therapeutic target.

LGR5

LGR5 is a 7-transmembrane domain receptor protein targeted by the Wnt signaling pathway, which has been found to mark the stem cell populations of the hair follicle in the skin, stomach and intestine. It has also been associated to cancer where it is overexpressed in multiple cancer types, including colorectal cancer, basal cell carcinomas and ovary and liver tumors (Schuijers and Clevers, 2012). The Wnt signaling pathway has also been shown to be essential to maintain pluripotency in ESCs (ten Berge et al., 2011).

PDGFRA

Platelet derived growth factor receptor alpha (PDGFRA) is a tyrosine kinase receptor required for embryonic development and is frequently activated in cancers where it promotes EMT; tumor growth, angiogenesis, invasion, and metastasis (Andrae et al., 2008). A recent report has shown that breast CSCs require a shift to PDGFRA signaling after activation of an EMT program (Tam et al., 2013).

PITX2

Pituitary homeobox-2 (*PITX2*) is a member of the bicoid/paired-like homeobox gene family, and is known to play a central role in determining left-right asymmetry in vertebrates and development of multiple organs. Here it serves as a downstream effector of Nodal, TGFβ, and the Wnt signaling pathway (Kioussi et al., 2002). It has also been shown to have an oncogenic role in thyroid (Huang et al., 2010).

Others

We also observed upregulation of a marker of an epidermal progenitor cell population, *Plet1*/MTS24 (Nijhof et al., 2006) **(Figure 8)**. Expansion of the epidermal stem cell compartments could also be related to tumorigenesis. Recently, a study has shown *Snail* overexpression in the basal layer of stratified epithelia can lead to proliferation of the keratinocyte stem cells shown by expansion of the $MTS24^{+}$ progenitor population and

ultimately this can also mediate the development of spontaneous skin tumors (De Craene et al., 2014).

In ESCs, NANOG is capable of binding to many sites and recruiting chromatin remodelers. Interestingly, one of our upregulated targets was *Hdac7*, a histone deacetylase involved in chromatin remodeling that could mediate aberrant gene silencing in tumors (Johnstone, 2002). In our case, ChIP of NANOG in ESCC cells showed direct binding to many promoters of upregulated genes in *Nanog*-overexpressing papillomas. NANOG has been shown to bind many genes in ESCs, by cooperating with SOX2 and OCT4 (Boyer et al., 2005; Kim et al., 2008). SOX2 has been shown to be an important tumor-promoting factor in SCC (Bass et al., 2009). Consistently, NANOG could also be cooperating with SOX2 in these tumor types, but a recent study of SOX2 binding partners in ESCC, did not show NANOG as one of the hits (Watanabe et al., 2014a). With ChIP-seq data of NANOG in ESCC we will be able to identify its transcriptional network, and compare it to its role in ESCs.

2.3. NANOG and microRNAs

MicroRNAs (miRNAs) have been shown to be key in somatic and pluripotent stem cells (Gangaraju and Lin, 2009) and are frequently deregulated in cancer (Zhang et al., 2006b). In cancer, miRNAs can act as tumor suppressors or oncogenes, depending on the cellular context (Spizzo et al., 2009). One of the miRNAs we found to be upregulated in *Nanog*-overexpressing back skin, papillomas and carcinomas was *miR-21*.

miR-21 has been extensively studied since it was found to be upregulated in cardiovascular diseases and cancer (Jazbutyte and Thum, 2010). *miR-21* has been demonstrated to be oncogenic, and it is also known as an oncomiR. Overexpression of miR-21 has a driving role in pre-B lymphoma and in Kras-mediated lung tumorigenesis (Hatley et al., 2010; Medina et al., 2010), and consistently, deletion of *miR-21* reduces tumorigenesis of skin carcinogenesis (DMA/TPA protocol), its oncogenic effect has been associated to upregulation of the ERK, AKT and JNK pathways (Ma et al., 2011). *miR-21* has also been associated to EMT and metastasis in SCCs (Bornachea et al., 2012). Previous studies have already linked NANOG to expression of miR-21 in HNSCC cells (Bourguignon et al., 2012a). In our case we have observed this *in vivo*, and we also confirmed that NANOG could directly bind to the *miR-21* promoter. This reinforces NANOG's oncogenic potential through upregulation of oncomiR miR-21.

What still are unknown are the regulators of NANOG activation in cancer. From our results, we can already see differences in ChIP performed in pluripotent NTERA2 cells compared to ESCC cells. In pluripotent cells, NANOG is known to bind its own promoter (Gifford et al., 2013), in our results we confirmed this finding in NTERA2 cells. On the other hand NANOG binding to its own promoter is absent in ESCC cells. This suggests that there are other regulators of *NANOG* expression in cancer or that its locus is epigentically inaccessible. Another possibility is that *NANOGP8* is being expressed in these cells and therefore follows a different regulation pattern compared to its paralog *NANOG1*. NANOG regulation has been extensively studied in ESCs, where numerous factors regulate its transcription (Apostolou et al., 2013), but this still should be further studied in the context of cancer.

We have demonstrated that NANOG has a driving role in tumor progression. To ultimately demonstrate if Nanog is a good target for cancer therapies, we would require addressing the requirement of NANOG during tumorigenesis. This would entail the use of a conditional KO *Nanog* strategy and similarly a chemical carcinogenesis approach to view if it possible to induce tumors (benign or malignant) after *Nanog* deletion.

All together, NANOG activation in cancer can lead to upregulation of multiple targets that are known to promote tumor progression and metastasis, which reinforces the oncogenic potential of NANOG in SCCs **(Figure 58)**.

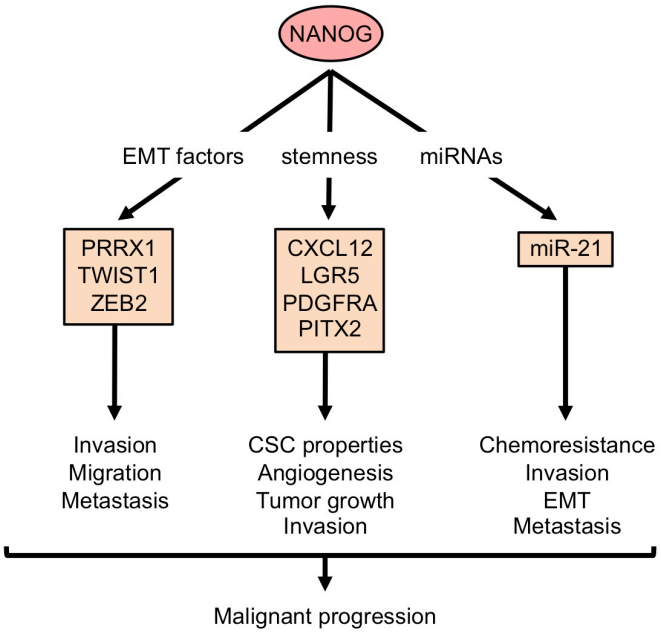


Figure 58. Oncogenic potential of NANOG in SCCs. We have observed NANOG can directly regulate EMT factors, *PRRX1, TWIST1* and *ZEB2*; stemness related factors *CXCL12, LGR5, PDGFRA* and *PITX2*; and oncomiR, *miR-21*. All these factors can contribute to NANOG's oncogenic potential in cancer.

3. NANOG variants

The *NANOG* gene family is highly complex in humans, but only three of its members have the capacity to produce full-length proteins, namely *NANOG1, NANOG2* and *NANOGP8*. *NANOG1* is highly expressed in pluripotent cells, where it is known to play a critical role and is considered one of the master regulators of pluripotency (Chambers et al., 2003; Chambers and Tomlinson, 2009; Mitsui et al., 2003; Wang et al., 2003). Although our current work is focused on *NANOGP8*, it is worth mentioning that we present evidence indicating that *NANOG2* is expressed in hESCs at significant levels.

NANOGP8 is gaining relevance since it has been shown to be expressed in many cancer cell lines and tumors where it contributes to oncogenic features, including CSC properties, proliferation, clonogenicity and tumorigenicity **(Table 1)**. Although, the protein encoded by *NANOGP8* only differs in two amino acids relative to *NANOG1*, these two amino acids are located in important domains for transcriptional regulation (Chang et al., 2009; Do et al., 2009; Pan and Pei, 2005), opening the possibility that *NANOGP8* function could differ from *NANOG1*. Also *NANOG1* and *NANOGP8* have distinct regulatory sequences, therefore activating factors of NANOG in cancer could be unrelated to its known regulators in ESCs.

With our results we demonstrate that *NANOGP8* is the functional variant most widely expressed in the analyzed cancer cell lines. At the protein level, NANOG1 and NANOGP8 are indistinguishable using antibodies due to their high degree of identity. With qRT-PCR, we discarded *NANOG1* expression in most cancer cell lines, confirming that NANOG protein arises from the *NANOGP8* gene. Distinctively, NANOG protein detection in cancer cells shows a complex mobility pattern in agreement with previous reports (Eberle et al., 2010; Jeter et al., 2009; Kim et al., 2005). The molecular bases are still unclear for this, but a number of potential mechanisms could account for the different NANOG mobilities, amongst them phosphorylation (Moretto-Zita et al., 2010), ubiquitylation (Moretto-Zita et al., 2010; Ramakrishna et al., 2011), and caspase-mediated cleavage (Fujita et al., 2008). With use of knockdown techniques, we confirm that the detected complex pattern corresponds to NANOG protein. Overexpression of NANOG variants in 293T cells also gives rise to different bands and the pattern differs among NANOG1, its point mutants and NANOGP8. This suggests different PTMs could be occurring in NANOG1 and NANOGP8, and also cellular context could be affecting this, as observed in the cancer cell lines.

We also provide evidence of the pluripotent capacity of NANOGP8. In particular, we have employed two assays, using human and mouse fibroblasts, and in both assays *NANOGP8*

improved reprogramming efficiency to the same extent as *NANOG1*. NANOG1 and NANOGP8 have an identical homeodomain, which could account for their equal effect in pluripotency. This is in agreement with other studies that have shown that the homeodomain is sufficient to establish pluripotency in *Nanog* deficient cells (Theunissen et al., 2011).

We have focused on NANOG variants with coding capacity, but it is also interesting that most cancer cell lines widely express high levels of pseudogenes NANOGP4 and NANOGP5. Although pseudogenes have always been considered to be biologically inconsequential due to a lack of an ORF, several reports have shown they may be relevant in regulation of their ancestral genes and to have a role in carcinogenesis (Poliseno et al., 2010). Further analysis should examine if these pseudogenes have an effect in expression of NANOG in cancer.

Our data demonstrate that NANOGP8 is also a reprogramming factor and, therefore, its widespread expression in cancer cells may contribute to the tumorigenic phenotype by promoting undifferentiated states, thus refueling cancer growth.

CONCLUSIONS

1. The pluripotency factor *Nanog* is expressed in adult tissues, specifically in the basal layer of stratified epithelia.

2. The ubiquitous overexpression of *Nanog* induces hyperplasia of mucosal stratified epithelial tissues, including the esophagus and forestomach, by induction of a mitotic transcriptional program, including direct binding to the *Aurka* promoter and its upregulation.

3. NANOG is expressed in esophageal squamous cell carcinomas (ESCC) and head and neck squamous cell carcinomas (HNSCC).

4. The inducible overexpression of *Nanog* in mouse skin epithelia favours the malignant conversion of skin tumors induced by chemical carcinogenesis, leading to increased SCC formation by upregulating epithelial-mesenchymal transition (EMT) and stemness drivers.

5. NANOG is a cell autonomous and direct activator of EMT and stemness programs in epithelial cells, including squamous cell carcinoma cells.

6. Human cancer cells show predominant expression of *NANOGP8*, whereas pluripotent cells (embryonic stem cells and embryonal carcinoma cells) express the *NANOG* variant *NANOG1*, and to a less extent, *NANOG2*.

7. *NANOGP8* possesses the same pluripotent activity as *NANOG1* in the reprogramming of human and mouse differentiated cells.

CONCLUSIONES

1. El factor de pluripotencia NANOG se expresa en tejidos adultos, específicamente en la capa basal de epitelio estatificado.

2. La sobreexpresión ubicua de *Nanog* induce hiperplasia de tejidos estratificados epiteliales mucosales al promover un programa transcripcional mitótico, incluyendo la regulación directa de la expresión de *Aurka* por unión a su promotor.

3. NANOG se expresa en carcinomas de células escamosas de esófago y de cara y cuello.

4. La sobreeexpresión de *Nanog* en la epidermis de ratón promueve la progresión maligna de tumores de piel inducidos por carcinogénesis epitelial, llevando a un incremento en la formación de carcinomas de células escamosas al sobre-regular genes clave implicados en transición epitelio-mesenquimal (EMT) y en troncalidad.

5. NANOG activa de forma directa programas de EMT y troncalidad en células epiteliales, incluyendo en células tumorales provenientes de carcinomas de células escamosas de cara y cuello.

6. Células tumorales humanas expresan de forma predominante la variante *NANOGP8*, mientras que células pluripotentes expresan la variante *NANOG1* y en menor medida *NANOG2*.

7. NANOGP8 posee la misma actividad pluripotente comparado a NANOG1 en la reprogramación de células diferenciadas de ratón y humano.

BIBLIOGRAPHY

Abad, M., Mosteiro, L., Pantoja, C., Canamero, M., Rayon, T., Ors, I., Grana, O., Megias, D., Dominguez, O., Martinez, D.*, et al.* (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340-345.

Abel, E.L., Angel, J.M., Kiguchi, K., and DiGiovanni, J. (2009). Multi-stage chemical carcinogenesis in mouse skin: fundamentals and applications. Nat Protoc *4*, 1350-1362.

Agrawal, N., Frederick, M.J., Pickering, C.R., Bettegowda, C., Chang, K., Li, R.J., Fakhry, C., Xie, T.X., Zhang, J., Wang, J.*, et al.* (2011). Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science *333*, 1154-1157.

Alam, M., and Ratner, D. (2001). Cutaneous squamous-cell carcinoma. N Engl J Med *344*, 975-983.

Alcolea, M.P., and Jones, P.H. (2014). Lineage analysis of epidermal stem cells. Cold Spring Harb Perspect Med *4*, a015206.

Ambady, S., Malcuit, C., Kashpur, O., Kole, D., Holmes, W.F., Hedblom, E., Page, R.L., and Dominko, T. (2010). Expression of NANOG and NANOGP8 in a variety of undifferentiated and differentiated human cells. Int J Dev Biol *54*, 1743-1754.

Andrae, J., Gallini, R., and Betsholtz, C. (2008). Role of platelet-derived growth factors in physiology and medicine. Genes Dev *22*, 1276-1312.

Andrews, P.W., Damjanov, I., Simon, D., Banting, G.S., Carlin, C., Dracopoli, N.C., and Fogh, J. (1984). Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro. Lab Invest *50*, 147-162.

Apostolou, E., Ferrari, F., Walsh, R.M., Bar-Nur, O., Stadtfeld, M., Cheloufi, S., Stuart, H.T., Polo, J.M., Ohsumi, T.K., Borowsky, M.L.*, et al.* (2013). Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. Cell Stem Cell *12*, 699-712.

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature *502*, 462-471.

Arama, E., Agapite, J., and Steller, H. (2003). Caspase activity and a specific cytochrome C are required for sperm differentiation in Drosophila. Dev Cell *4*, 687-697.

Arnold, I., and Watt, F.M. (2001). c-Myc activation in transgenic mouse epidermis results in mobilization of stem cells and differentiation of their progeny. Curr Biol *11*, 558-568.

Arnold, K., Sarkar, A., Yram, M.A., Polo, J.M., Bronson, R., Sengupta, S., Seandel, M., Geijsen, N., and Hochedlinger, K. (2011). Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell *9*, 317-329.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol *10*, 91-103.

Arwert, E.N., Hoste, E., and Watt, F.M. (2012). Epithelial stem cells, wound healing and cancer. Nat Rev Cancer *12*, 170-180.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Backvall, H., Asplund, A., Gustafsson, A., Sivertsson, A., Lundeberg, J., and Ponten, F. (2005). Genetic tumor archeology: microdissection and genetic heterogeneity in squamous and basal cell carcinoma. Mutat Res *571*, 65-79.

Badeaux, M.A., Jeter, C.R., Gong, S., Liu, B., Suraneni, M.V., Rundhaug, J., Fischer, S.M., Yang, T., Kusewitt, D., and Tang, D.G. (2013). In vivo functional studies of tumor-specific retrogene NanogP8 in transgenic animals. Cell Cycle *12*, 2395-2408.

Baggiolini, M. (1998). Chemokines and leukocyte traffic. Nature *392*, 565-568.

Bahl, K., Saraya, A., and Sharma, R. (2012). Increased Levels of Circulating and Tissue mRNAs of Oct-4, Sox-2, Bmi-1 and Nanog is ESCC Patients: Potential Tool for Minimally Invasive Cancer Diagnosis. Biomark Insights *7*, 27-37.

Balkwill, F. (2004). Cancer and the chemokine network. Nat Rev Cancer *4*, 540-550.

Balmain, A., Ramsden, M., Bowden, G.T., and Smith, J. (1984). Activation of the mouse cellular Harvey-ras gene in chemically induced benign skin papillomas. Nature *307*, 658-660.

Barker, N., Bartfeld, S., and Clevers, H. (2010). Tissue-resident adult stem cell populations of rapidly self-renewing organs. Cell Stem Cell *7*, 656-670.

Bass, A.J., Watanabe, H., Mermel, C.H., Yu, S., Perner, S., Verhaak, R.G., Kim, S.Y., Wardwell, L., Tamayo, P., Gat-Viks, I.*, et al.* (2009). SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet *41*, 1238-1242.

Berdnik, D., and Knoblich, J.A. (2002). Drosophila Aurora-A is required for centrosome maturation and actin-dependent asymmetric protein localization during mitosis. Curr Biol *12*, 640-647.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Biernaskie, J., Paris, M., Morozova, O., Fagan, B.M., Marra, M., Pevny, L., and Miller, F.D. (2009). SKPs derive from hair follicle precursors and exhibit properties of adult dermal stem cells. Cell Stem Cell *5*, 610-623.

Bischoff, J.R., Anderson, L., Zhu, Y., Mossie, K., Ng, L., Souza, B., Schryver, B., Flanagan, P., Clairvoyant, F., Ginther, C.*, et al.* (1998). A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. EMBO J *17*, 3052-3065.

Blanpain, C., and Fuchs, E. (2009). Epidermal homeostasis: a balancing act of stem cells in the skin. Nat Rev Mol Cell Biol *10*, 207-217.

Blanpain, C., Horsley, V., and Fuchs, E. (2007). Epithelial stem cells: turning over new leaves. Cell *128*, 445-458.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martinez, J.A., Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A.*, et al.* (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature *500*, 222-226.

Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006). Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. Stem Cells *24*, 2007-2013.

Blick, T., Hugo, H., Widodo, E., Waltham, M., Pinto, C., Mani, S.A., Weinberg, R.A., Neve, R.M., Lenburg, M.E., and Thompson, E.W. (2010). Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer. J Mammary Gland Biol Neoplasia *15*, 235-252.

Booth, H.A., and Holland, P.W. (2004). Eleven daughters of NANOG. Genomics *84*, 229-238.

Bornachea, O., Santos, M., Martinez-Cruz, A.B., Garcia-Escudero, R., Duenas, M., Costa, C., Segrelles, C., Lorz, C., Buitrago, A., Saiz-Ladera, C.*, et al.* (2012). EMT and induction of miR-21 mediate metastasis development in Trp53-deficient tumours. Sci Rep *2*, 434.

Bourguignon, L.Y., Earle, C., Wong, G., Spevak, C.C., and Krueger, K. (2012a). Stem cell marker (Nanog) and Stat-3 signaling promote MicroRNA-21 expression and chemoresistance in hyaluronan/CD44-activated head and neck squamous cell carcinoma cells. Oncogene *31*, 149-160.

Bourguignon, L.Y., Peyrollier, K., Xia, W., and Gilad, E. (2008). Hyaluronan-CD44 interaction activates stem cell marker Nanog, Stat-3-mediated MDR1 gene expression, and ankyrin-regulated multidrug efflux in breast and ovarian tumor cells. J Biol Chem *283*, 17635-17651.

Bourguignon, L.Y., Spevak, C.C., Wong, G., Xia, W., and Gilad, E. (2009). Hyaluronan-CD44 interaction with protein kinase C(epsilon) promotes oncogenic signaling by the stem cell marker Nanog and the Production of microRNA-21, leading to down-regulation of the tumor suppressor protein PDCD4, anti-apoptosis, and chemotherapy resistance in breast tumor cells. J Biol Chem *284*, 26533-26546.

Bourguignon, L.Y., Wong, G., Earle, C., and Chen, L. (2012b). Hyaluronan-CD44v3 interaction with Oct4-Sox2-Nanog promotes miR-302 expression leading to self-renewal, clonal formation, and cisplatin resistance in cancer stem cells from head and neck squamous cell carcinoma. J Biol Chem *287*, 32800-32824.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev Biol *227*, 239-255.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. Nat Rev Cancer *5*, 744-749.

Brash, D.E., Rudolph, J.A., Simon, J.A., Lin, A., McKenna, G.J., Baden, H.P., Halperin, A.J., and Ponten, J. (1991). A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc Natl Acad Sci U S A *88*, 10124-10128.

Brattain, M.G., Fine, W.D., Khaled, F.M., Thompson, J., and Brattain, D.E. (1981). Heterogeneity of malignant cells from a human colonic carcinoma. Cancer Res *41*, 1751-1756.

Brown, K., Strathdee, D., Bryson, S., Lambie, W., and Balmain, A. (1998). The malignant capacity of skin tumours induced by expression of a mutant H-ras transgene depends on the cell type targeted. Curr Biol *8*, 516-524.

Brumbaugh, J., Russell, J.D., Yu, P., Westphall, M.S., Coon, J.J., and Thomson, J.A. (2014). NANOG Is Multiply Phosphorylated and Directly Modified by ERK2 and CDK1 In Vitro. Stem Cell Reports *2*, 18-25.

Bubenik, J., Baresova, M., Viklicky, V., Jakoubkova, J., Sainerova, H., and Donner, J. (1973). Established cell line of urinary bladder carcinoma (T24) containing tumour-specific antigen. Int J Cancer *11*, 765-773.

Buchmann, A., Ruggeri, B., Klein-Szanto, A.J., and Balmain, A. (1991). Progression of squamous carcinoma cells to spindle carcinomas of mouse skin is associated with an imbalance of H-ras alleles on chromosome 7. Cancer Res *51*, 4097-4101.

Burdon, T., Smith, A., and Savatier, P. (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol *12*, 432-438.

Burger, J.A., and Kipps, T.J. (2006). CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. Blood *107*, 1761-1767.

Burns, P.A., Kemp, C.J., Gannon, J.V., Lane, D.P., Bremner, R., and Balmain, A. (1991). Loss of heterozygosity and mutational alterations of the p53 gene in skin tumours of interspecific hybrid mice. Oncogene *6*, 2363-2369.

Cai, N., Li, M., Qu, J., Liu, G.H., and Izpisua Belmonte, J.C. (2012). Post-translational modulation of pluripotency. J Mol Cell Biol *4*, 262-265.

Cailleau, R., Young, R., Olive, M., and Reeves, W.J., Jr. (1974). Breast tumor cell lines from pleural effusions. J Natl Cancer Inst *53*, 661-674.

Camp, E., Sanchez-Sanchez, A.V., Garcia-Espana, A., Desalle, R., Odqvist, L., Enrique O'Connor, J., and Mullor, J.L. (2009). Nanog regulates proliferation during early fish development. Stem Cells *27*, 2081-2091.

Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol *2*, 76-83.

Canon, S., Herranz, C., and Manzanares, M. (2006). Germ cell restricted expression of chick Nanog. Dev Dyn *235*, 2889-2894.

Carmena, M., and Earnshaw, W.C. (2003). The cellular geography of aurora kinases. Nat Rev Mol Cell Biol *4*, 842-854.

Castellanos, E., Dominguez, P., and Gonzalez, C. (2008). Centrosome dysfunction in Drosophila neural stem cells causes tumors that are not due to genome instability. Curr Biol *18*, 1209-1214.

Caussinus, E., and Gonzalez, C. (2005). Induction of tumor growth by altered stem-cell asymmetric division in Drosophila melanogaster. Nat Genet *37*, 1125-1129.

Cavallaro, M., Mariani, J., Lancini, C., Latorre, E., Caccia, R., Gullo, F., Valotta, M., DeBiasi, S., Spinardi, L., Ronchi, A.*, et al.* (2008). Impaired generation of mature neurons by neural stem cells from hypomorphic Sox2 mutants. Development *135*, 541-557.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

Chambers, I., and Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene *23*, 7150-7160.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311-2322.

Chang, C.C., Hsu, W.H., Wang, C.C., Chou, C.H., Kuo, M.Y., Lin, B.R., Chen, S.T., Tai, S.K., Kuo, M.L., and Yang, M.H. (2013). Connective tissue growth factor activates pluripotency genes and mesenchymal-epithelial transition in head and neck cancer cells. Cancer Res *73*, 4147-4157.

Chang, D.F., Tsai, S.C., Wang, X.C., Xia, P., Senadheera, D., and Lutzko, C. (2009). Molecular characterization of the human NANOG protein. Stem Cells *27*, 812-821.

Chatterjee, A., Upadhyay, S., Chang, X., Nagpal, J.K., Trink, B., and Sidransky, D. (2008). U-box-type ubiquitin E4 ligase, UFD2a attenuates cisplatin mediated degradation of DeltaNp63alpha. Cell Cycle *7*, 1231-1237.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell *10*, 615-624.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J.*, et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Chen, Y.C., Chen, Y.W., Hsu, H.S., Tseng, L.M., Huang, P.I., Lu, K.H., Chen, D.T., Tai, L.K., Yung, M.C., Chang, S.C.*, et al.* (2009). Aldehyde dehydrogenase 1 is a putative marker for cancer stem cells in head and neck squamous cancer. Biochem Biophys Res Commun *385*, 307-313.

Chen, Y.S., Huang, W.L., Chang, S.H., Chang, K.W., Kao, S.Y., Lo, J.F., and Su, P.F. (2013). Enhanced filopodium formation and stem-like phenotypes in a novel metastatic head and neck cancer cell model. Oncol Rep *30*, 2829-2837.

Chiou, S.H., Wang, M.L., Chou, Y.T., Chen, C.J., Hong, C.F., Hsieh, W.J., Chang, H.T., Chen, Y.S., Lin, T.W., Hsu, H.S.*, et al.* (2010). Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. Cancer Res *70*, 10433-10444.

Chiou, S.H., Yu, C.C., Huang, C.Y., Lin, S.C., Liu, C.J., Tsai, T.H., Chou, S.H., Chien, C.S., Ku, H.H., and Lo, J.F. (2008). Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. Clin Cancer Res *14*, 4085-4095.

Chute, J.P., Muramoto, G.G., Whitesides, J., Colvin, M., Safi, R., Chao, N.J., and McDonnell, D.P. (2006). Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. Proc Natl Acad Sci U S A *103*, 11707-11712.

Corominas, M., Sloan, S.R., Leon, J., Kamino, H., Newcomb, E.W., and Pellicer, A. (1991). ras activation in human tumors and in animal model systems. Environ Health Perspect *93*, 19-25.

Cufi, S., Bonavia, R., Vazquez-Martin, A., Oliveras-Ferraros, C., Corominas-Faja, B., Cuyas, E., Martin-Castillo, B., Barrajon-Catalan, E., Visa, J., Segura-Carretero, A.*, et al.* (2013). Silibinin suppresses EMT-driven erlotinib resistance by reversing the high miR-21/low miR-200c signature in vivo. Sci Rep *3*, 2459.

Cui, W., Fowlis, D.J., Bryson, S., Duffie, E., Ireland, H., Balmain, A., and Akhurst, R.J. (1996). TGFbeta1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. Cell *86*, 531-542.

D'Souza, G., Kreimer, A.R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W.M., Westra, W.H., and Gillison, M.L. (2007). Case-control study of human papillomavirus and oropharyngeal cancer. N Engl J Med *356*, 1944-1956.

Dar, A.A., Belkhiri, A., and El-Rifai, W. (2009). The aurora kinase A regulates GSK-3beta in gastric cancer cells. Oncogene *28*, 866-875.

Das, S., Jena, S., and Levasseur, D.N. (2011). Alternative splicing produces Nanog protein variants with different capacities for self-renewal and pluripotency in embryonic stem cells. J Biol Chem *286*, 42690-42703.

De Craene, B., Denecker, G., Vermassen, P., Taminau, J., Mauch, C., Derore, A., Jonkers, J., Fuchs, E., and Berx, G. (2014). Epidermal Snail expression drives skin cancer initiation and progression through enhanced cytoprotection, epidermal stem/progenitor cell expansion and enhanced metastatic potential. Cell Death Differ *21*, 310-320.

Dixon, J.E., Allegrucci, C., Redwood, C., Kump, K., Bian, Y., Chatfield, J., Chen, Y.H., Sottile, V., Voss, S.R., Alberio, R.*, et al.* (2010). Axolotl Nanog activity in mouse embryonic stem cells

demonstrates that ground state pluripotency is conserved from urodele amphibians to mammals. Development *137*, 2973-2980.

Do, H.J., Lee, W.Y., Lim, H.Y., Oh, J.H., Kim, D.K., Kim, J.H., and Kim, T. (2009). Two potent transactivation domains in the C-terminal region of human NANOG mediate transcriptional activation in human embryonic carcinoma cells. J Cell Biochem *106*, 1079-1089.

Do, H.J., Lim, H.Y., Kim, J.H., Song, H., and Chung, H.M. (2007). An intact homeobox domain is required for complete nuclear localization of human Nanog. Biochem Biophys Res Commun *353*, 770-775.

Doitsidou, M., Reichman-Fried, M., Stebler, J., Koprunner, M., Dorries, J., Meyer, D., Esguerra, C.V., Leung, T., and Raz, E. (2002). Guidance of primordial germ cell migration by the chemokine SDF-1. Cell *111*, 647-659.

Doupe, D.P., Alcolea, M.P., Roshan, A., Zhang, G., Klein, A.M., Simons, B.D., and Jones, P.H. (2012). A single progenitor population switches behavior to maintain and repair esophageal epithelium. Science *337*, 1091-1093.

Driskell, R.R., Clavel, C., Rendl, M., and Watt, F.M. (2011). Hair follicle dermal papilla cells at a glance. J Cell Sci *124*, 1179-1182.

Driskell, R.R., Giangreco, A., Jensen, K.B., Mulder, K.W., and Watt, F.M. (2009). Sox2-positive dermal papilla cells specify hair follicle type in mammalian epidermis. Development *136*, 2815-2823.

Du, Y., Shi, L., Wang, T., Liu, Z., and Wang, Z. (2012). Nanog siRNA plus Cisplatin may enhance the sensitivity of chemotherapy in esophageal cancer. J Cancer Res Clin Oncol *138*, 1759-1767.

Duda, D.G., Kozin, S.V., Kirkpatrick, N.D., Xu, L., Fukumura, D., and Jain, R.K. (2011). CXCL12 (SDF1alpha)-CXCR4/CXCR7 pathway inhibition: an emerging sensitizer for anticancer therapies? Clin Cancer Res *17*, 2074-2080.

Dykxhoorn, D.M. (2010). MicroRNAs and metastasis: little RNAs go a long way. Cancer Res *70*, 6401-6406.

Earnshaw, W.C., Martins, L.M., and Kaufmann, S.H. (1999). Mammalian caspases: structure, activation, substrates, and functions during apoptosis. Annu Rev Biochem *68*, 383-424.

Easty, D.M., Easty, G.C., Carter, R.L., Monaghan, P., and Butler, L.J. (1981). Ten human carcinoma cell lines derived from squamous carcinomas of the head and neck. Br J Cancer *43*, 772-785.

Eberle, I., Pless, B., Braun, M., Dingermann, T., and Marschalek, R. (2010). Transcriptional properties of human NANOG1 and NANOG2 in acute leukemic cells. Nucleic Acids Res *38*, 5384-5395.

Ellis, P., Fagan, B.M., Magness, S.T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M., and Pevny, L. (2004). SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. Dev Neurosci *26*, 148-165.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Faddah, D.A., Wang, H., Cheng, A.W., Katz, Y., Buganim, Y., and Jaenisch, R. (2013). Single-cell analysis reveals that expression of nanog is biallelic and equally variable as that of other pluripotency factors in mouse ESCs. Cell Stem Cell *13*, 23-29.

Falkner, F.G., and Zachau, H.G. (1984). Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. Nature *310*, 71-74.

Favaro, R., Valotta, M., Ferri, A.L., Latorre, E., Mariani, J., Giachino, C., Lancini, C., Tosetti, V., Ottolenghi, S., Taylor, V.*, et al.* (2009). Hippocampal development and neural stem cell maintenance require Sox2-dependent regulation of Shh. Nat Neurosci *12*, 1248-1256.

Feinstein-Rotkopf, Y., and Arama, E. (2009). Can't live without them, can live with them: roles of caspases during vital cellular processes. Apoptosis *14*, 980-995.

Feng, Y., Broder, C.C., Kennedy, P.E., and Berger, E.A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science *272*, 872-877.

Ferri, A.L., Cavallaro, M., Braida, D., Di Cristofano, A., Canta, A., Vezzani, A., Ottolenghi, S., Pandolfi, P.P., Sala, M., DeBiasi, S.*, et al.* (2004). Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. Development *131*, 3805-3819.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell *11*, 477-490.

Festuccia, N., Osorno, R., Wilson, V., and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. Curr Opin Genet Dev *23*, 504-511.

Filipczyk, A., Gkatzis, K., Fu, J., Hoppe, P.S., Lickert, H., Anastassiadis, K., and Schroeder, T. (2013). Biallelic expression of nanog protein in mouse embryonic stem cells. Cell Stem Cell *13*, 12-13.

Friedmann-Morvinski, D., and Verma, I.M. (2014). Dedifferentiation and reprogramming: origins of cancer stem cells. EMBO Rep *15*, 244-253.

Fuchs, E. (2009). Finding one's niche in the skin. Cell Stem Cell *4*, 499-502.

Fujita, J., Crane, A.M., Souza, M.K., Dejosez, M., Kyba, M., Flavell, R.A., Thomson, J.A., and Zwaka, T.P. (2008). Caspase activity mediates the differentiation of embryonic stem cells. Cell Stem Cell *2*, 595-601.

Fujiwara, D., Kato, K., Nohara, S., Iwanuma, Y., and Kajiyama, Y. (2013). The usefulness of three-dimensional cell culture in induction of cancer stem cells from esophageal squamous cell carcinoma cell lines. Biochem Biophys Res Commun *434*, 773-778.

Fusenig, N.E., Amer, S.M., Boukamp, P., and Worst, P.K. (1978). Characteristics of chemically transformed mouse epidermal cells in vitro and in vivo. Bull Cancer *65*, 271-279.

Gandarillas, A., and Watt, F.M. (1997). c-Myc promotes differentiation of human epidermal stem cells. Genes Dev *11*, 2869-2882.

Gangaraju, V.K., and Lin, H. (2009). MicroRNAs: key regulators of stem cells. Nat Rev Mol Cell Biol *10*, 116-125.

Garzon, R., Marcucci, G., and Croce, C.M. (2010). Targeting microRNAs in cancer: rationale, strategies and challenges. Nat Rev Drug Discov *9*, 775-789.

Gat, U., DasGupta, R., Degenstein, L., and Fuchs, E. (1998). De Novo hair follicle morphogenesis and hair tumors in mice expressing a truncated beta-catenin in skin. Cell *95*, 605-614.

Gazdar, A.F., Carney, D.N., Nau, M.M., and Minna, J.D. (1985). Characterization of variant subclasses of cell lines derived from small cell lung cancer having distinctive biochemical, morphological, and growth properties. Cancer Res *45*, 2924-2930.

Gazdar, A.F., Carney, D.N., Russell, E.K., Sims, H.L., Baylin, S.B., Bunn, P.A., Jr., Guccion, J.G., and Minna, J.D. (1980). Establishment of continuous, clonable cultures of small-cell

carcinoma of lung which have amine precursor uptake and decarboxylation cell properties. Cancer Res *40*, 3502-3507.

Gazdar, A.F., Helman, L.J., Israel, M.A., Russell, E.K., Linnoila, R.I., Mulshine, J.L., Schuller, H.M., and Park, J.G. (1988). Expression of neuroendocrine cell markers L-dopa decarboxylase, chromogranin A, and dense core granules in human tumors of endocrine and nonendocrine origin. Cancer Res *48*, 4078-4082.

Gehring, W.J., Qian, Y.Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A.F., Resendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994). Homeodomain-DNA recognition. Cell *78*, 211-223.

Gey, G., Coffman, W., and Kubicek, M.T. (1952). Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. Cancer research *12*, 264-265.

Giard, D.J., Aaronson, S.A., Todaro, G.J., Arnstein, P., Kersey, J.H., Dosik, H., and Parks, W.P. (1973). In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J Natl Cancer Inst *51*, 1417-1423.

Gidekel, S., Pizov, G., Bergman, Y., and Pikarsky, E. (2003). Oct-3/4 is a dose-dependent oncogenic fate determinant. Cancer Cell *4*, 361-370.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R.*, et al.* (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153*, 1149-1163.

Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S.*, et al.* (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell *1*, 555-567.

Gioanni, J., Fischel, J.L., Lambert, J.C., Demard, F., Mazeau, C., Zanghellini, E., Ettore, F., Formento, P., Chauvel, P., Lalanne, C.M.*, et al.* (1988). Two new human tumor cell lines derived from squamous cell carcinomas of the tongue: establishment, characterization and response to cytotoxic treatment. Eur J Cancer Clin Oncol *24*, 1445-1455.

Goon, P.K., Stanley, M.A., Ebmeyer, J., Steinstrasser, L., Upile, T., Jerjes, W., Bernal-Sprekelsen, M., Gorner, M., and Sudhoff, H.H. (2009). HPV & head and neck cancer: a descriptive update. Head Neck Oncol *1*, 1-8.

Grachtchouk, M., Mo, R., Yu, S., Zhang, X., Sasaki, H., Hui, C.C., and Dlugosz, A.A. (2000). Basal cell carcinomas in mice overexpressing Gli2 in skin. Nat Genet *24*, 216-217.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Silencing of core transcription factors in human EC cells highlights the importance of autocrine FGF signaling for self-renewal. BMC Dev Biol *7*, 46.

Grenman, R., Carey, T.E., McClatchey, K.D., Wagner, J.G., Pekkola-Heino, K., Schwartz, D.R., Wolf, G.T., Lacivita, L.P., Ho, L., Baker, S.R.*, et al.* (1991). In vitro radiation resistance among cell lines established from patients with squamous cell carcinoma of the head and neck. Cancer *67*, 2741-2747.

Grigoriadis, A.E., Kennedy, M., Bozec, A., Brunton, F., Stenbeck, G., Park, I.H., Wagner, E.F., and Keller, G.M. (2010). Directed differentiation of hematopoietic precursors and functional osteoclasts from human ES and iPS cells. Blood *115*, 2769-2776.

Griner, E.M., and Kazanietz, M.G. (2007). Protein kinase C and other diacylglycerol effectors in cancer. Nat Rev Cancer *7*, 281-294.

Gruschus, J.M., Tsao, D.H., Wang, L.H., Nirenberg, M., and Ferretti, J.A. (1997). Interactions of the vnd/NK-2 homeodomain with DNA by nuclear magnetic resonance spectroscopy: basis of binding specificity. Biochemistry *36*, 5372-5380.

Gupta, P.B., Onder, T.T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R.A., and Lander, E.S. (2009). Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell *138*, 645-659.

Han, J., Zhang, F., Yu, M., Zhao, P., Ji, W., Zhang, H., Wu, B., Wang, Y., and Niu, R. (2012). RNA interference-mediated silencing of NANOG reduces cell proliferation and induces G0/G1 cell cycle arrest in breast cancer cells. Cancer Lett *321*, 80-88.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creyghton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. Nature *462*, 595-601.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508-525.

Harvey, R.P. (1996). NK-2 homeobox genes and heart development. Dev Biol *178*, 203-216.

Hasmim, M., Noman, M.Z., Messai, Y., Bordereaux, D., Gros, G., Baud, V., and Chouaib, S. (2013). Cutting edge: Hypoxia-induced Nanog favors the intratumoral infiltration of regulatory T cells and macrophages via direct regulation of TGF-beta1. J Immunol *191*, 5802-5806.

Hatley, M.E., Patrick, D.M., Garcia, M.R., Richardson, J.A., Bassel-Duby, R., van Rooij, E., and Olson, E.N. (2010). Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. Cancer Cell *18*, 282-293.

Hattori, N., Imao, Y., Nishino, K., Ohgane, J., Yagi, S., Tanaka, S., and Shiota, K. (2007). Epigenetic regulation of Nanog gene in embryonic stem and trophoblast stem cells. Genes Cells *12*, 387-396.

Havran, W.L., and Jameson, J.M. (2010). Epidermal T cells and wound healing. J Immunol *184*, 5423-5428.

Hennings, H., Glick, A.B., Lowry, D.T., Krsmanovic, L.S., Sly, L.M., and Yuspa, S.H. (1993). FVB/N mice: an inbred strain sensitive to the chemical induction of squamous cell carcinomas in the skin. Carcinogenesis *14*, 2353-2358.

Hermann, P.C., Huber, S.L., Herrler, T., Aicher, A., Ellwart, J.W., Guba, M., Bruns, C.J., and Heeschen, C. (2007). Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. Cell Stem Cell *1*, 313-323.

Ho, B., Olson, G., Figel, S., Gelman, I., Cance, W.G., and Golubovskaya, V.M. (2012). Nanog increases focal adhesion kinase (FAK) promoter activity and expression and directly binds to FAK protein to be phosphorylated. J Biol Chem *287*, 18656-18673.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465-477.

Horoszewicz, J.S., Leong, S.S., Kawinski, E., Karr, J.P., Rosenthal, H., Chu, T.M., Mirand, E.A., and Murphy, G.P. (1983). LNCaP model of human prostatic carcinoma. Cancer Res *43*, 1809-1818.

Huang, Y., Guigon, C.J., Fan, J., Cheng, S.Y., and Zhu, G.Z. (2010). Pituitary homeobox 2 (PITX2) promotes thyroid carcinogenesis by activation of cyclin D2. Cell Cycle *9*, 1333-1341.

Ibrahim, E.E., Babaei-Jadidi, R., Saadeddin, A., Spencer-Dene, B., Hossaini, S., Abuzinadah, M., Li, N., Fadhil, W., Ilyas, M., Bonnet, D.*, et al.* (2012). Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. Stem Cells *30*, 2076-2087.

Imamura, M., Miura, K., Iwabuchi, K., Ichisaka, T., Nakagawa, M., Lee, J., Kanatsu-Shinohara, M., Shinohara, T., and Yamanaka, S. (2006). Transcriptional repression and DNA

hypermethylation of a small set of ES cell marker genes in male germline stem cells. BMC Dev Biol *6*, 34.

Ishiguro, T., Sato, A., Ohata, H., Sakai, H., Nakagama, H., and Okamoto, K. (2012). Differential expression of nanog1 and nanogp8 in colon cancer cells. Biochem Biophys Res Commun *418*, 199-204.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533-538.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Curr Biol *4*, 1-7.

Janes, S.M., and Watt, F.M. (2006). New roles for integrins in squamous-cell carcinoma. Nat Rev Cancer *6*, 175-183.

Jauch, R., Ng, C.K., Saikatendu, K.S., Stevens, R.C., and Kolatkar, P.R. (2008). Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. J Mol Biol *376*, 758-770.

Jazbutyte, V., and Thum, T. (2010). MicroRNA-21: from cancer to cardiovascular disease. Curr Drug Targets *11*, 926-935.

Jechlinger, M., Grunert, S., Tamir, I.H., Janda, E., Ludemann, S., Waerner, T., Seither, P., Weith, A., Beug, H., and Kraut, N. (2003). Expression profiling of epithelial plasticity in tumor progression. Oncogene *22*, 7155-7169.

Jeter, C.R., Badeaux, M., Choy, G., Chandra, D., Patrawala, L., Liu, C., Calhoun-Davis, T., Zaehres, H., Daley, G.Q., and Tang, D.G. (2009). Functional evidence that the self-renewal gene NANOG regulates human tumor development. Stem Cells *27*, 993-1005.

Jeter, C.R., Liu, B., Liu, X., Chen, X., Liu, C., Calhoun-Davis, T., Repass, J., Zaehres, H., Shen, J.J., and Tang, D.G. (2011). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. Oncogene *30*, 3833-3845.

Johnstone, R.W. (2002). Histone-deacetylase inhibitors: novel drugs for the treatment of cancer. Nat Rev Drug Discov *1*, 287-299.

Juuri, E., Saito, K., Ahtiainen, L., Seidel, K., Tummers, M., Hochedlinger, K., Klein, O.D., Thesleff, I., and Michon, F. (2012). Sox2+ stem cells contribute to all epithelial lineages of the tooth via Sfrp5+ progenitors. Dev Cell *23*, 317-328.

Kaighn, M.E., Narayan, K.S., Ohnuki, Y., Lechner, J.F., and Jones, L.W. (1979). Establishment and characterization of a human prostatic carcinoma cell line (PC-3). Invest Urol *17*, 16-23.

Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. Cell *39*, 499-509.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M.*, et al.* (2014). The UCSC Genome Browser database: 2014 update. Nucleic Acids Res *42*, D764-770.

Kemp, C.J., Donehower, L.A., Bradley, A., and Balmain, A. (1993). Reduction of p53 gene dosage does not increase initiation or promotion but enhances malignant progression of chemically induced skin tumors. Cell *74*, 813-822.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Kim, J.S., Kim, J., Kim, B.S., Chung, H.Y., Lee, Y.Y., Park, C.S., Lee, Y.S., Lee, Y.H., and Chung, I.Y. (2005). Identification and functional characterization of an alternative splice variant within the fourth exon of human nanog. Exp Mol Med *37*, 601-607.

Kioussi, C., Briata, P., Baek, S.H., Rose, D.W., Hamblet, N.S., Herman, T., Ohgi, K.A., Lin, C., Gleiberman, A., Wang, J.*, et al.* (2002). Identification of a Wnt/Dvl/beta-Catenin --> Pitx2 pathway mediating cell-type-specific proliferation during development. Cell *111*, 673-685.

Klein-Szanto, A.J. (1989). Pathology of human and experimental skin tumors. Carcinog Compr Surv *11*, 19-53.

Klemm, J.D., and Pabo, C.O. (1996). Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. Genes Dev *10*, 27-36.

Knowles, B.B., Howe, C.C., and Aden, D.P. (1980). Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. Science *209*, 497-499.

Komori, H.K., Meehan, T.F., and Havran, W.L. (2006). Epithelial and mucosal gamma delta T cells. Curr Opin Immunol *18*, 534-538.

Krause, C.J., Carey, T.E., Ott, R.W., Hurbis, C., McClatchey, K.D., and Regezi, J.A. (1981). Human squamous cell carcinoma. Establishment and characterization of new permanent cell lines. Arch Otolaryngol *107*, 703-710.

Kulesz-Martin, M., Kilkenny, A.E., Holbrook, K.A., Digernes, V., and Yuspa, S.H. (1983). Properties of carcinogen altered mouse epidermal cells resistant to calcium-induced terminal differentiation. Carcinogenesis *4*, 1367-1377.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lapouge, G., Youssef, K.K., Vokaer, B., Achouri, Y., Michaux, C., Sotiropoulou, P.A., and Blanpain, C. (2011). Identifying the cellular origin of squamous skin tumors. Proc Natl Acad Sci U S A *108*, 7431-7436.

Lecker, S.H., Goldberg, A.L., and Mitch, W.E. (2006). Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. J Am Soc Nephrol *17*, 1807-1819.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R., Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-renewal. Cell Stem Cell *1*, 403-415.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liedtke, S., Enczmann, J., Waclawczyk, S., Wernet, P., and Kogler, G. (2007). Oct4 and its pseudogenes confuse stem cell research. Cell Stem Cell *1*, 364-366.

Lin, C.W., Liao, M.Y., Lin, W.W., Wang, Y.P., Lu, T.Y., and Wu, H.C. (2012). Epithelial cell adhesion molecule regulates tumor initiation and tumorigenesis via activating reprogramming factors and epithelial-mesenchymal transition gene expression in colon cancer. J Biol Chem *287*, 39449-39459.

Lin, C.W., Lin, J.C., and Prout, G.R., Jr. (1985). Establishment and characterization of four human bladder tumor cell lines and sublines with different degrees of malignancy. Cancer Res *45*, 5070-5079.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat Cell Biol *7*, 165-171.

Liu, B., Kissinger, C.R., and Pabo, C.O. (1990). Crystallization and preliminary X-ray diffraction studies of the engrailed homeodomain and of an engrailed homeodomain/DNA complex. Biochem Biophys Res Commun *171*, 257-259.

Liu, K., Jiang, M., Lu, Y., Chen, H., Sun, J., Wu, S., Ku, W.Y., Nakagawa, H., Kita, Y., Natsugoe, S.*, et al.* (2013). Sox2 cooperates with inflammation-mediated Stat3 activation in the malignant transformation of foregut basal progenitor cells. Cell Stem Cell *12*, 304-315.

Lobo, N.A., Shimono, Y., Qian, D., and Clarke, M.F. (2007). The biology of cancer stem cells. Annu Rev Cell Dev Biol *23*, 675-699.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Lomas, A., Leonardi-Bee, J., and Bath-Hextall, F. (2012). A systematic review of worldwide incidence of nonmelanoma skin cancer. Br J Dermatol *166*, 1069-1080.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci U S A *105*, 2883-2888.

Lu, L.Y., Wood, J.L., Ye, L., Minter-Dykhouse, K., Saunders, T.L., Yu, X., and Chen, J. (2008). Aurora A is essential for early embryonic development and tumor suppression. J Biol Chem *283*, 31785-31790.

Lu, X., Mazur, S.J., Lin, T., Appella, E., and Xu, Y. (2013). The pluripotency factor nanog promotes breast cancer tumorigenesis and metastasis. Oncogene *Oncogene advance online publication, 17 June 2013*.

Ma, X., Kumar, M., Choudhury, S.N., Becker Buscaglia, L.E., Barker, J.R., Kanakamedala, K., Liu, M.F., and Li, Y. (2011). Loss of the miR-21 allele elevates the expression of its target genes and reduces tumorigenesis. Proc Natl Acad Sci U S A *108*, 10144-10149.

Macleod, A.S., and Havran, W.L. (2011). Functions of skin-resident gammadelta T cells. Cell Mol Life Sci *68*, 2399-2408.

Madan, V., Lear, J.T., and Szeimies, R.M. (2010). Non-melanoma skin cancer. Lancet *375*, 673-685.

Magni, M., Shammah, S., Schiro, R., Mellado, W., Dalla-Favera, R., and Gianni, A.M. (1996). Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer. Blood *87*, 1097-1103.

Mallo, M., Schrewe, H., Martin, J.F., Olson, E.N., and Ohnemus, S. (2000). Assembling a functional tympanic membrane: signals from the external acoustic meatus coordinate development of the malleal manubrium. Development *127*, 4127-4136.

Malumbres, M. (2011). Physiological relevance of cell cycle kinases. Physiol Rev *91*, 973-1007.

Manfredi, M.G., Ecsedy, J.A., Chakravarty, A., Silverman, L., Zhang, M., Hoar, K.M., Stroud, S.G., Chen, W., Shinde, V., Huck, J.J.*, et al.* (2011). Characterization of Alisertib (MLN8237), an investigational small-molecule inhibitor of aurora A kinase using novel in vivo pharmacodynamic assays. Clin Cancer Res *17*, 7614-7624.

Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M.*, et al.* (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell *133*, 704-715.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A.*, et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Medema, J.P. (2013). Cancer stem cells: the challenges ahead. Nat Cell Biol *15*, 338-344.

Medina, P.P., Nolde, M., and Slack, F.J. (2010). OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature *467*, 86-90.

Mehra, R., Serebriiskii, I.G., Burtness, B., Astsaturov, I., and Golemis, E.A. (2013). Aurora kinases in head and neck cancer. Lancet Oncol *14*, e425-435.

Meng, H.M., Zheng, P., Wang, X.Y., Liu, C., Sui, H.M., Wu, S.J., Zhou, J., Ding, Y.Q., and Li, J.M. (2010). Overexpression of nanog predicts tumor progression and poor prognosis in colorectal cancer. Cancer Biol Ther *9*, 295-302.

Meraldi, P., Honda, R., and Nigg, E.A. (2002). Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53-/- cells. EMBO J *21*, 483-492.

Messerschmidt, D.M., and Kemler, R. (2010). Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. Dev Biol *344*, 129-137.

Mills, A.A., Zheng, B., Wang, X.J., Vogel, H., Roop, D.R., and Bradley, A. (1999). p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature *398*, 708-713.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470-473.

Molinolo, A.A., Amornphimoltham, P., Squarize, C.H., Castilho, R.M., Patel, V., and Gutkind, J.S. (2009). Dysregulated molecular networks in head and neck carcinogenesis. Oral Oncol *45*, 324-334.

Moretto-Zita, M., Jin, H., Shen, Z., Zhao, T., Briggs, S.P., and Xu, Y. (2010). Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. Proc Natl Acad Sci U S A *107*, 13312-13317.

Morris, R.J. (2000). Keratinocyte stem cells: targets for cutaneous carcinogens. J Clin Invest *106*, 3-8.

Morris, R.J., Coulter, K., Tryson, K., and Steinberg, S.R. (1997). Evidence that cutaneous carcinogen-initiated epithelial cells from mice are quiescent rather than actively cycling. Cancer Res *57*, 3436-3443.

Morris, R.J., Fischer, S.M., and Slaga, T.J. (1986). Evidence that a slowly cycling subpopulation of adult murine epidermal cells retains carcinogen. Cancer Res *46*, 3061-3066.

Mullin, N.P., Yates, A., Rowe, A.J., Nijmeijer, B., Colby, D., Barlow, P.N., Walkinshaw, M.D., and Chambers, I. (2008). The pluripotency rheostat Nanog functions as a dimer. Biochem J *411*, 227-231.

Navarro, J.M., Casatorres, J., and Jorcano, J.L. (1995). Elements controlling the expression and induction of the skin hyperproliferation-associated keratin K6. J Biol Chem *270*, 21362-21367.

Neumuller, R.A., and Knoblich, J.A. (2009). Dividing cellular asymmetry: asymmetric cell division and its implications for stem cells and cancer. Genes Dev *23*, 2675-2699.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487-492.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Niessen, M.T., Iden, S., and Niessen, C.M. (2012). The in vivo function of mammalian cell and tissue polarity regulators--how to shape and maintain the epidermal barrier. J Cell Sci *125*, 3501-3510.

Nieto, M.A. (2013). Epithelial plasticity: a common theme in embryonic and cancer cells. Science *342*, 1234850.

Nieto, M.A., and Cano, A. (2012). The epithelial-mesenchymal transition under control: global programs to regulate epithelial plasticity. Semin Cancer Biol *22*, 361-368.

Nijhof, J.G., Braun, K.M., Giangreco, A., van Pelt, C., Kawamoto, H., Boyd, R.L., Willemze, R., Mullenders, L.H., Watt, F.M., de Gruijl, F.R.*, et al.* (2006). The cell-surface marker MTS24 identifies a novel population of follicular keratinocytes with characteristics of progenitor cells. Development *133*, 3027-3037.

Nilsson, M., Unden, A.B., Krause, D., Malmqwist, U., Raza, K., Zaphiropoulos, P.G., and Toftgard, R. (2000). Induction of basal cell carcinomas and trichoepitheliomas in mice overexpressing GLI-1. Proc Natl Acad Sci U S A *97*, 3438-3443.

Nishihira, T., Hashimoto, Y., Katayama, M., Mori, S., and Kuroki, T. (1993). Molecular and cellular features of esophageal cancer cells. J Cancer Res Clin Oncol *119*, 441-449.

Nishihira, T., Kasai, M., Mori, S., Watanabe, T., Kuriya, Y., Suda, M., Kitamura, M., Hirayama, K., Akaishi, T., and Sasaki, T. (1979). Characteristics of two cell lines (TE-1 and TE-2) derived from human squamous cell carcinoma of the esophagus. Gann *70*, 575-584.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev *12*, 2048-2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118-122.

Noh, K.H., Kim, B.W., Song, K.H., Cho, H., Lee, Y.H., Kim, J.H., Chung, J.Y., Hewitt, S.M., Seong, S.Y., Mao, C.P.*, et al.* (2012). Nanog signaling in cancer promotes stem-like phenotype and immune evasion. J Clin Invest *122*, 4077-4093.

Ocana, O.H., Corcoles, R., Fabra, A., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., and Nieto, M.A. (2012). Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell *22*, 709-724.

Oft, M., Akhurst, R.J., and Balmain, A. (2002). Metastasis is driven by sequential elevation of H-ras and Smad2 levels. Nat Cell Biol *4*, 487-494.

Oh, J.H., Do, H.J., Yang, H.M., Moon, S.Y., Cha, K.Y., Chung, H.M., and Kim, J.H. (2005). Identification of a putative transactivation domain in human Nanog. Exp Mol Med *37*, 250-254.

Okano, J., Gaslightwala, I., Birnbaum, M.J., Rustgi, A.K., and Nakagawa, H. (2000). Akt/protein kinase B isoforms are differentially regulated by epidermal growth factor stimulation. J Biol Chem *275*, 30934-30942.

Okubo, T., Clark, C., and Hogan, B.L. (2009). Cell lineage mapping of taste bud cells and keratinocytes in the mouse tongue and soft palate. Stem Cells *27*, 442-450.

Palmero, I., and Serrano, M. (2001). Induction of senescence by oncogenic Ras. Methods Enzymol *333*, 247-256.

Pan, G., and Pei, D. (2005). The stem cell pluripotency factor NANOG activates transcription with two unusually potent subdomains at its C terminus. J Biol Chem *280*, 1401-1407.

Pan, G., and Thomson, J.A. (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. Cell Res *17*, 42-49.

Pan, G.J., and Pei, D.Q. (2003). Identification of two distinct transactivation domains in the pluripotency sustaining factor nanog. Cell Res *13*, 499-502.

Parslow, T.G., Blair, D.L., Murphy, W.J., and Granner, D.K. (1984). Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. Proc Natl Acad Sci U S A *81*, 2650-2654.

Pennathur, A., Gibson, M.K., Jobe, B.A., and Luketich, J.D. (2013). Oesophageal carcinoma. Lancet *381*, 400-412.

Perez de Castro, I., Aguirre-Portoles, C., Fernandez-Miranda, G., Canamero, M., Cowley, D.O., Van Dyke, T., and Malumbres, M. (2013). Requirements for Aurora-A in tissue regeneration and tumor development in adult mammals. Cancer Res *73*, 6804-6815.

Perez-Losada, J., and Balmain, A. (2003). Stem-cell hierarchy in skin cancer. Nat Rev Cancer *3*, 434-443.

Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. J Mol Biol *302*, 1023-1039.

Po, A., Ferretti, E., Miele, E., De Smaele, E., Paganelli, A., Canettieri, G., Coni, S., Di Marcotullio, L., Biffoni, M., Massimi, L.*, et al.* (2010). Hedgehog controls neural stem cells through p53-independent regulation of Nanog. EMBO J *29*, 2646-2658.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033-1038.

Ponten, J., and Saksela, E. (1967). Two established in vitro cell lines from human mesenchymal tumours. Int J Cancer *2*, 434-447.

Que, J., Luo, X., Schwartz, R.J., and Hogan, B.L. (2009). Multiple roles for Sox2 in the developing and adult mouse trachea. Development *136*, 1899-1907.

Quigley, D.A., To, M.D., Perez-Losada, J., Pelorosso, F.G., Mao, J.H., Nagase, H., Ginzinger, D.G., and Balmain, A. (2009). Genetic architecture of mouse skin inflammation and tumour susceptibility. Nature *458*, 505-508.

Quintanilla, M., Brown, K., Ramsden, M., and Balmain, A. (1986). Carcinogen-specific mutation and amplification of Ha-ras during mouse skin carcinogenesis. Nature *322*, 78-80.

Quintanilla, M., Haddow, S., Jonas, D., Jaffe, D., Bowden, G.T., and Balmain, A. (1991). Comparison of ras activation during epidermal carcinogenesis in vitro and in vivo. Carcinogenesis *12*, 1875-1881.

Ramakrishna, S., Suresh, B., Lim, K.H., Cha, B.H., Lee, S.H., Kim, K.S., and Baek, K.H. (2011). PEST motif sequence regulating human NANOG for proteasomal degradation. Stem Cells Dev *20*, 1511-1519.

Regan, J.L., Sourisseau, T., Soady, K., Kendrick, H., McCarthy, A., Tang, C., Brennan, K., Linardopoulos, S., White, D.E., and Smalley, M.J. (2013). Aurora A kinase regulates mammary epithelial cell fate by determining mitotic spindle orientation in a Notch-dependent manner. Cell Rep *4*, 110-123.

Robertson, M., Stenhouse, F., Colby, D., Marland, J.R., Nichols, J., Tweedie, S., and Chambers, I. (2006). Nanog retrotransposed genes with functionally conserved open reading frames. Mamm Genome *17*, 732-743.

Rocco, J.W., Leong, C.O., Kuperwasser, N., DeYoung, M.P., and Ellisen, L.W. (2006). p63 mediates survival in squamous cell carcinoma by suppression of p73-dependent apoptosis. Cancer Cell *9*, 45-56.

Romagnani, P., Lasagni, L., Annunziato, F., Serio, M., and Romagnani, S. (2004). CXC chemokines: the regulatory link between inflammation and angiogenesis. Trends Immunol *25*, 201-209.

Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G.*, et al.* (2013). ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res *41*, D56-63.

Rudin, C.M., Durinck, S., Stawiski, E.W., Poirier, J.T., Modrusan, Z., Shames, D.S., Bergbower, E.A., Guan, Y., Shin, J., Guillory, J.*, et al.* (2012). Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet *44*, 1111-1116.

Rundhaug, J.E., and Fischer, S.M. (2010). Molecular mechanisms of mouse skin tumor promotion. Cancers (Basel) *2*, 436-482.

Sanchez-Sanchez, A.V., Camp, E., Leal-Tassias, A., Atkinson, S.P., Armstrong, L., Diaz-Llopis, M., and Mullor, J.L. (2010). Nanog regulates primordial germ cell migration through Cxcr4b. Stem Cells *28*, 1457-1464.

Sanchez-Tillo, E., Liu, Y., de Barrios, O., Siles, L., Fanlo, L., Cuatrecasas, M., Darling, D.S., Dean, D.C., Castells, A., and Postigo, A. (2012). EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness. Cell Mol Life Sci *69*, 3429-3456.

Sarkar, A., and Hochedlinger, K. (2013). The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell *12*, 15-30.

Sasai, K., Parant, J.M., Brandt, M.E., Carter, J., Adams, H.P., Stass, S.A., Killary, A.M., Katayama, H., and Sen, S. (2008). Targeted disruption of Aurora A causes abnormal mitotic spindle assembly, chromosome misalignment and embryonic lethality. Oncogene *27*, 4122-4127.

Saunders, A., Faiola, F., and Wang, J. (2013). Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. Stem Cells *31*, 1227-1236.

Schuff, M., Siegel, D., Philipp, M., Bundschu, K., Heymann, N., Donow, C., and Knochel, W. (2012). Characterization of Danio rerio Nanog and functional comparison to Xenopus Vents. Stem Cells Dev *21*, 1225-1238.

Schuijers, J., and Clevers, H. (2012). Adult mammalian stem cells: the role of Wnt, Lgr5 and R-spondins. EMBO J *31*, 2685-2696.

Schwarz, B.A., Bar-Nur, O., Silva, J.C., and Hochedlinger, K. (2014). Nanog is dispensable for the generation of induced pluripotent stem cells. Curr Biol *24*, 347-350.

Shan, J., Shen, J., Liu, L., Xia, F., Xu, C., Duan, G., Xu, Y., Ma, Q., Yang, Z., Zhang, Q.*, et al.* (2012). Nanog regulates self-renewal of cancer stem cells through the insulin-like growth factor pathway in human hepatocellular carcinoma. Hepatology *56*, 1004-1014.

Shigeishi, H., Biddle, A., Gammon, L., Emich, H., Rodini, C.O., Gemenetzidis, E., Fazil, B., Sugiyama, M., Kamata, N., and Mackenzie, I.C. (2013). Maintenance of stem cell self-renewal in head and neck cancers requires actions of GSK3beta influenced by CD44 and RHAMM. Stem Cells *31*, 2073-2083.

Shimada, Y., Okumura, T., Sekine, S., Moriyama, M., Sawada, S., Matsui, K., Yoshioka, I., Hojo, S., Yoshida, T., Nagata, T., *et al.* (2012). Expression analysis of iPS cell - inductive genes in esophageal squamous cell carcinoma by tissue microarray. Anticancer Res *32*, 5507-5514.

Siegel, R., DeSantis, C., Virgo, K., Stein, K., Mariotto, A., Smith, T., Cooper, D., Gansler, T., Lerro, C., Fedewa, S., *et al.* (2012). Cancer treatment and survivorship statistics, 2012. CA Cancer J Clin *62*, 220-241.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol *6*, e253.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722-737.

Simons, B.D., and Clevers, H. (2011). Strategies for homeostatic stem cell self-renewal in adult tissues. Cell *145*, 851-862.

Siu, M.K., Wong, E.S., Kong, D.S., Chan, H.Y., Jiang, L., Wong, O.G., Lam, E.W., Chan, K.K., Ngan, H.Y., Le, X.F., *et al.* (2013). Stem cell transcription factor NANOG controls cell migration and invasion via dysregulation of E-cadherin and FoxJ1 and contributes to adverse clinical outcome in ovarian cancers. Oncogene *32*, 3500-3509.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.

Soule, H.D., Vazguez, J., Long, A., Albert, S., and Brennan, M. (1973). A human cell line from a pleural effusion derived from a breast carcinoma. J Natl Cancer Inst *51*, 1409-1416.

Spencer, J.M., Kahn, S.M., Jiang, W., DeLeo, V.A., and Weinstein, I.B. (1995). Activated ras genes occur in human actinic keratoses, premalignant precursors to squamous cell carcinomas. Arch Dermatol *131*, 796-800.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O., and Thomson, J.A. (2003). Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc Natl Acad Sci U S A *100*, 13350-13355.

Spizzo, R., Nicoloso, M.S., Croce, C.M., and Calin, G.A. (2009). SnapShot: MicroRNAs in Cancer. Cell *137*, 586-586 e581.

Sramkoski, R.M., Pretlow, T.G., 2nd, Giaconia, J.M., Pretlow, T.P., Schwartz, S., Sy, M.S., Marengo, S.R., Rhim, J.S., Zhang, D., and Jacobberger, J.W. (1999). A new human prostate carcinoma cell line, 22Rv1. In Vitro Cell Dev Biol Anim *35*, 403-409.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364-377.

Stone, K.R., Mickey, D.D., Wunderli, H., Mickey, G.H., and Paulson, D.F. (1978). Isolation of a human prostate carcinoma cell line (DU 145). Int J Cancer *21*, 274-281.

Stransky, N., Egloff, A.M., Tward, A.D., Kostic, A.D., Cibulskis, K., Sivachenko, A., Kryukov, G.V., Lawrence, M.S., Sougnez, C., McKenna, A., *et al.* (2011). The mutational landscape of head and neck squamous cell carcinoma. Science *333*, 1157-1160.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Suh, H., Consiglio, A., Ray, J., Sawai, T., D'Amour, K.A., and Gage, F.H. (2007). In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. Cell Stem Cell *1*, 515-528.

Sun, C., Sun, L., Jiang, K., Gao, D.M., Kang, X.N., Wang, C., Zhang, S., Huang, S., Qin, X., Li, Y.*, et al.* (2013). NANOG promotes liver cancer cell invasion by inducing epithelial-mesenchymal transition through NODAL/SMAD3 signaling pathway. Int J Biochem Cell Biol *45*, 1099-1108.

Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. Cell *128*, 747-762.

Suva, M.L., Riggi, N., and Bernstein, B.E. (2013). Epigenetic reprogramming in cancer. Science *339*, 1567-1570.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tam, W.L., Lu, H., Buikhuisen, J., Soh, B.S., Lim, E., Reinhardt, F., Wu, Z.J., Krall, J.A., Bierie, B., Guo, W.*, et al.* (2013). Protein kinase C alpha is a central signaling node and therapeutic target for breast cancer stem cells. Cancer Cell *24*, 347-364.

Tanaka, S., Kamachi, Y., Tanouchi, A., Hamada, H., Jing, N., and Kondoh, H. (2004). Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. Mol Cell Biol *24*, 8834-8846.

Tanaka, Y., Era, T., Nishikawa, S., and Kawamata, S. (2007). Forced expression of Nanog in hematopoietic stem cells results in a gammadeltaT-cell disorder. Blood *110*, 107-115.

Taranova, O.V., Magness, S.T., Fagan, B.M., Wu, Y., Surzenko, N., Hutton, S.R., and Pevny, L.H. (2006). SOX2 is a dose-dependent regulator of retinal neural progenitor competence. Genes Dev *20*, 1187-1202.

Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y., Gupta, S., Yang, J., Hartwell, K., Onder, T.T., Gupta, P.B., Evans, K.W.*, et al.* (2010). Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc Natl Acad Sci U S A *107*, 15449-15454.

Teicher, B.A., and Fricker, S.P. (2010). CXCL12 (SDF-1)/CXCR4 pathway in cancer. Clin Cancer Res *16*, 2927-2931.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol *13*, 1070-1075.

Theunissen, T.W., van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., and Silva, J.C. (2011). Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. Curr Biol *21*, 65-71.

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871-890.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell *22*, 725-736.

Tsunematsu, T., Takihara, Y., Ishimaru, N., Pagano, M., Takata, T., and Kudo, Y. (2013). Aurora-A controls pre-replicative complex assembly and DNA replication by stabilizing geminin in mitosis. Nat Commun *4*, 1885.

Uchino, K., Hirano, G., Hirahashi, M., Isobe, T., Shirakawa, T., Kusaba, H., Baba, E., Tsuneyoshi, M., and Akashi, K. (2012). Human Nanog pseudogene8 promotes the proliferation of gastrointestinal cancer cells. Exp Cell Res *318*, 1799-1807.

van der Schroeff, J.G., Evers, L.M., Boot, A.J., and Bos, J.L. (1990). Ras oncogene mutations in basal cell carcinomas and squamous cell carcinomas of human skin. J Invest Dermatol *94*, 423-425.

Vitale-Cross, L., Amornphimoltham, P., Fisher, G., Molinolo, A.A., and Gutkind, J.S. (2004). Conditional expression of K-ras in an epithelial compartment that includes the stem cells is sufficient to promote squamous cell carcinogenesis. Cancer Res *64*, 8804-8807.

Waikel, R.L., Kawachi, Y., Waikel, P.A., Wang, X.J., and Roop, D.R. (2001). Deregulated expression of c-Myc depletes epidermal stem cells. Nat Genet *28*, 165-168.

Wang, H., Somers, G.W., Bashirullah, A., Heberlein, U., Yu, F., and Chia, W. (2006a). Aurora-A acts as a tumor suppressor and regulates self-renewal of Drosophila neuroblasts. Genes Dev *20*, 3453-3463.

Wang, J., Levasseur, D.N., and Orkin, S.H. (2008). Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. Proc Natl Acad Sci U S A *105*, 6326-6331.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006b). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.

Wang, S.H., Tsai, M.S., Chiang, M.F., and Li, H. (2003). A novel NK-type homeobox gene, ENK (early embryo specific NK), preferentially expressed in embryonic stem cells. Gene Expr Patterns *3*, 99-103.

Watanabe, H., Ma, Q., Peng, S., Adelmant, G., Swain, D., Song, W., Fox, C., Francis, J.M., Pedamallu, C.S., Deluca, D.S., *et al.* (2014a). SOX2 and p63 colocalize at genetic loci in squamous cell carcinomas. J Clin Invest.

Watanabe, M., Ohnishi, Y., Inoue, H., Wato, M., Tanaka, A., Kakudo, K., and Nozaki, M. (2014b). NANOG expression correlates with differentiation, metastasis and resistance to preoperative adjuvant therapy in oral squamous cell carcinoma. Oncol Lett *7*, 35-40.

Welling, M., and Geijsen, N. (2013). Uncovering the true identity of naive pluripotent stem cells. Trends Cell Biol *23*, 442-448.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

White, A.C., Tran, K., Khuu, J., Dang, C., Cui, Y., Binder, S.W., and Lowry, W.E. (2011). Defining the origins of Ras/p53-mediated squamous cell carcinoma. Proc Natl Acad Sci U S A *108*, 7425-7430.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

SUPPLEMENTARY MATERIAL

Table S1. Differentially expressed genes in NANOG-overexpressing papillomas versus controls (DEG) (red: EMT-related genes; blue: stemness-related genes).

DEG (FDR<0.05)[a]
Upregulated in TG (n=3) vs CTR (n=4) papillomas

GENE	Log2 (fold change)	NOM p-val	FDR q-val
Atp2a1	6,55967	0	0
Ttn	7,82508	4,44089E-16	2,48312E-12
Ckm	6,68657	1,76119E-11	4,92385E-08
Car3	7,56808	1,46252E-11	4,92385E-08
Myh1	6,19826	2,97422E-11	6,65214E-08
Rps4y2	4,83277	3,1267E-10	5,82765E-07
Myh4	9,76944	3,86907E-10	6,18112E-07
Neb	4,78852	1,05064E-09	1,30548E-06
Nanog	3,3504	2,89841E-09	2,86892E-06
H19	5,14505	3,33506E-09	2,86892E-06
Pygm	5,47646	3,15503E-09	2,86892E-06
Myh2	6,82759	2,9939E-09	2,86892E-06
Fhl1	3,27325	5,39424E-09	4,30885E-06
Cxcl5	4,01	7,31023E-09	5,45002E-06
Ldb3	5,1954	9,29079E-09	6,49368E-06
Tspan1	6,34179	1,34298E-08	8,83443E-06
Myl1	7,41907	1,82127E-08	1,13151E-05
Hdac7	0,507551	2,88639E-08	1,69887E-05
Muc4	5,43117	4,65719E-08	2,60407E-05
Slc6a14	3,3864	5,37199E-08	2,86071E-05
Gm9573	4,1772	6,145E-08	2,98781E-05
Mybpc2	6,61558	5,92222E-08	2,98781E-05
Ppbp	4,45723	6,57444E-08	3,06341E-05
Xirp2	8,84208	8,05238E-08	3,60199E-05
Cmya5	7,71368	9,43437E-08	4,05787E-05
Plet1	3,23664	1,44838E-07	5,99897E-05
Liph	0,824535	1,87919E-07	7,50535E-05
Dynap	4,28125	2,90942E-07	0,000108454
Acta1	3,42611	4,84392E-07	0,00017474
Cxcl12/Sdf-1	3,48094	8,7566E-07	0,000306016
Mylpf	3,604	1,18551E-06	0,000401745
Tnnc2	7,86144	1,66602E-06	0,000547974
Slfn4	3,15787	2,49068E-06	0,000773701
Des	3,83497	2,43019E-06	0,000773701
Stfa2l1	2,52101	2,64669E-06	0,000799944
Hand1	4,7656	3,03465E-06	0,000893066
Slc15a2	5,59142	5,62084E-06	0,00157145
Ppp1r3a	7,49906	7,86352E-06	0,00214482
Actn3	7,91031	9,18878E-06	0,00244662
Acpp	1,34887	1,01814E-05	0,00260871
Prnd	1,69616	1,07306E-05	0,00260871
Slc5a8	2,13379	1,06291E-05	0,00260871
Scn8a	2,81525	1,13129E-05	0,00269175
Cxcl3	2,82839	1,15923E-05	0,00270075
C3	3,62271	1,46215E-05	0,00333699
Thbs4	5,83514	1,62731E-05	0,00363964
Stfa2	2,51483	2,27001E-05	0,00489649
Gprc5a	2,98638	2,27682E-05	0,00489649
Ivl	2,38136	2,42188E-05	0,00501553
Aldh1a3	2,13392	2,63686E-05	0,00536145
Myom1	3,71599	2,92021E-05	0,00583156
Myh8	6,86279	3,32031E-05	0,00651421
Col14a1	3,37758	4,04415E-05	0,00779753

Nov	4,20451	4,12846E-05	0,00782518
Irg1	3,22614	4,80172E-05	0,00894961
Tgm2	2,91676	5,20011E-05	0,00953325
Nmes1	2,58447	6,08827E-05	0,0109815
Mb	5,11013	6,29326E-05	0,011032
Cyp2g1	6,25924	7,84593E-05	0,0132941
Ryr1	3,14644	8,07453E-05	0,0134772
Cilp	2,90297	8,34107E-05	0,0137174
Casq1	5,88079	8,68645E-05	0,0140783
Atp1a2	3,84829	9,05208E-05	0,0144613
Lcn2	2,36315	0,000104649	0,016254
Sell	3,91241	0,000107728	0,016503
Plat	2,21234	0,000110605	0,0167148
Clec4e	3,35531	0,000115642	0,0172429
Myoz1	5,24086	0,000135532	0,0196838
Jph2	2,56037	0,000147204	0,0211049
Lcp1	1,75582	0,000155746	0,022047
Krt71	4,64116	0,000199048	0,0274809
Sec14l2	2,09105	0,000211102	0,0284428
Slco3a1	1,13427	0,000221865	0,0288502
Scara5	4,58217	0,000221295	0,0288502
Vcan	1,02296	0,000232356	0,029625
Cpn1	2,95868	0,000235771	0,029625
Myom2	6,00486	0,000235284	0,029625
Gm5416	2,20526	0,000247856	0,0304691
Cd300lf	3,29543	0,000252813	0,0307305
Fa2h	4,0605	0,000266412	0,0320353
Ltf	3,26081	0,000271416	0,0322898
Sgcg	6,2053	0,000281906	0,0331848
Pkia	3,97673	0,000300045	0,0349521
Retnla	3,374	0,000333794	0,0384827
Plau	1,58562	0,000337397	0,0385011
Lgals1	1,77197	0,000343726	0,0388272
Il1b	2,24167	0,00036124	0,0403974
Srgn	2,15106	0,000375829	0,0416128
Tmc5	3,77452	0,00038531	0,0422443
Mybpc1	1,7977E+308	0,000393614	0,0427358
Tcap	4,36964	0,000434202	0,0465468
Prrx1	1,39363	0,000443961	0,0468379
Wnt5a	1,31548	0,000449316	0,0469598
Hdc	2,7227	0,000467022	0,0483584
Tdh	1,80571	0,000476752	0,048913
Lpo	2,40731	0,000485072	0,0493142
Nrap	6,51767	0,000495117	0,0498819
Tcap	4,36964	0,000434202	0,0465468

[a] FDR q-val, false discovery rate q-value; NOM p-val, nominal p-value

Published articles directly related to this PhD thesis:

1. Palla, A.R., Piazzolla, D., Abad M., Li, H., Dominguez, O., Schönthaler, H.B., Wagner, E.F. and Serrano, M. (2013) Reprogramming activity of NANOGP8, a NANOG family member widely expressed in cancer. *Oncogene advance online publication 10 June 2013; doi: 10.1038/onc.2013.196*

2. Palla, A.R., Piazzolla, D., Cañamero, M., Graña, O., Gómez-López, G., Dueñas, M., Paramio, J.M. and Serrano, M. (2014) The pluripotency factor NANOG promotes squamous cell carcinoma. *Submitted*.

3. Piazzolla, D., Palla, A.R., Pantoja, C., Cañamero, M., Pérez-Castro, I., Ortega, S., Gómez-López, G., Dominguez, O., Megías, D., Roncador, G., Luque-Garcia, J.L., Fernandez-Tresguerres, B., Fernandez, A.F., Fraga, M.F., Rodríguez-Justo, M., Manzanares, M., Sánchez-Carbayo, M., García-Pedrero, J.M.; Rodrigo, J.P., Malumbres, M., and Serrano, M. (2014). Lineage-restricted function of the pluripotency factor NANOG in stratified epithelia. *Nature Communications*, in revision.